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

Short title: Phylogeny and diversification of sea urchins

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 much of their early history uncertain. The origin of sand dollars, one of its most distinctive clades, is also unclear due to an unstable phylogenetic context and discrepancies between molecular divergence times and fossil evidence. We employ seventeen novel genomes and transcriptomes to build a phylogenomic dataset with a near-complete sampling of major lineages. With it, we revise the phylogeny and divergence times of echinoids, and place their history within the broader context of echinoderm evolution. We also 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 the use of site-heterogeneous models shows little effects. The choice of loci shows an intermediate impact, affecting mostly deep Paleozoic nodes, for which clock-like genes recover dates more congruent with fossil evidence. Our results reveal that crown group echinoids originated in the Permian and diversified rapidly in the Triassic, despite the relative lack of fossil evidence for this early diversification. We also clarify the relationships among sand dollars and their close relatives, showing that the genus *Apatopygus* represents a relict lineage with a deep Jurassic origin. Surprisingly, the origin of sand dollars is confidently dated to the Cretaceous, implying ghost ranges spanning approximately 50 million years, a remarkable discrepancy with their rich fossil record.

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

Introduction

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

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

Compared to the morphological conservatism of regular sea urchins, the evolutionary history of the relatively younger Irregularia was characterized by dramatic levels of morphological and ecological innovation [21-24]. Within the diversity of irregulars, sand dollars are the most easily recognized (Fig. 1). The clade includes greatly flattened forms that live in high-energy sandy environments where they feed using a unique mechanism for selecting and transporting organic particles to the mouth, where these are crushed using well-developed jaws [25, 26]. Sand dollars (Scutelloida) were long thought to be most closely related to sea biscuits (Clypeasteroida) given a wealth of shared morphological characters [19, 25]. The extraordinary fossil record of both sand dollars and sea biscuits suggested their last common ancestor originated in the early Cenozoic from among an assemblage known as “cassiduloids” [23, 25], a once diverse group that is today represented by three depauperate lineages: cassidulids (and close relatives), echinolampadids, and apatopygids [19, 27]. These taxa not only lack the defining features of both scutelloids and clypeasteroids but have experienced little morphological change since their origin deep in the Mesozoic [24, 27-29]. However, early molecular phylogenies supported both 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 (Fig. 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 (Fig. 1; [25, 31, 32]).

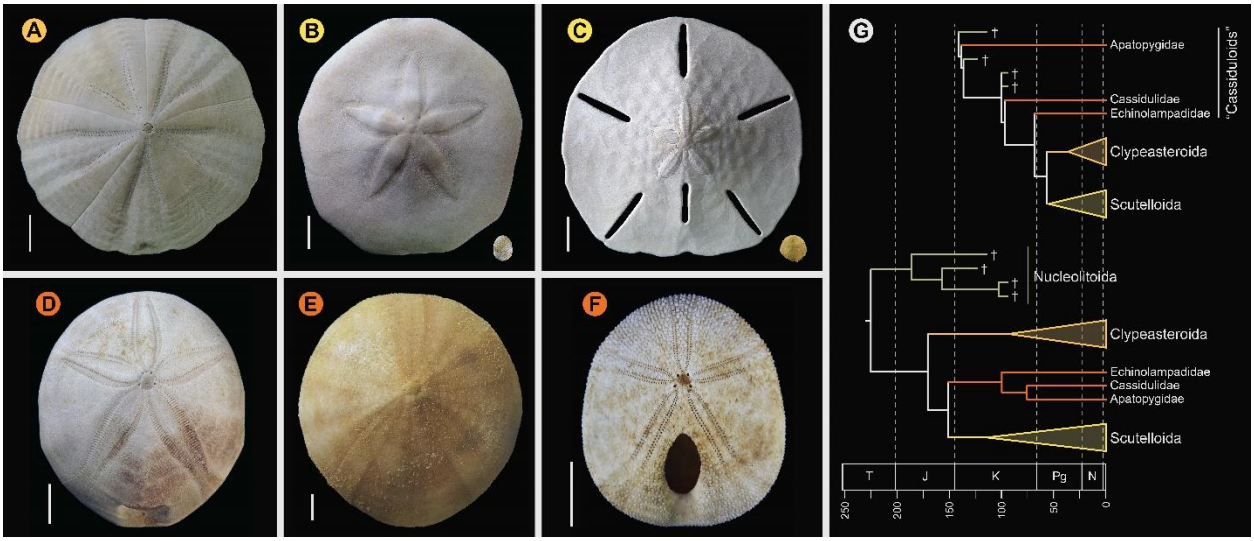


Fig. 1: Neognathostomate diversity and phylogenetic relationships. **A.** *Fellaster zelandiae*, North Island, New Zealand (Clypeasteroida). **B.** Large specimen: *Peronella japonica*, Ryukyu Islands, Japan; Small specimen: *Echinocyamus crispus*, Maricaban Island, Philippines (Laganina: Scutelloida).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

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

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

131

132 **Results**

133 Phylogenetic relationships supported by the full dataset were remarkably stable, with all nodes
134 but one being identically resolved and fully supported across a wide range of inference methods (Fig.
135 2A). While recovering a topology similar to those of previous molecular studies [3, 4, 13, 14, 30, 36], this
136 analysis is the first to sample and confidently place micropygoids and aspidodiadematoids within
137 Aulodonta, as well as resolve the relationships among all major clades of Neognathostomata
138 (scutelloids, clypeateroids and the three lineages of extant "cassiduloids"). Our results show that
139 *Apatopygus recens* is not related to the remaining "cassiduloids" but is instead the sister clade to all
140 other sampled neognathostomates. The strong support for this placement, as well as for a clade of
141 cassidulids and echinolampadids (Cassiduloida *sensu stricto*) as the sister group to sand dollars, provides
142 a basis for an otherwise elusive phylogenetic classification of neognathostomates. Our topology also
143 confirms that *Sinaechinocyamus mai*, a miniaturized species once considered a plesiomorphic member
144 of Scutelloida based on the reduction or loss of diagnostic features (Fig. 1), is in fact a derived
145 paedomorphic lineage closely related to *Scaphechinus mirabilis* [32].

146

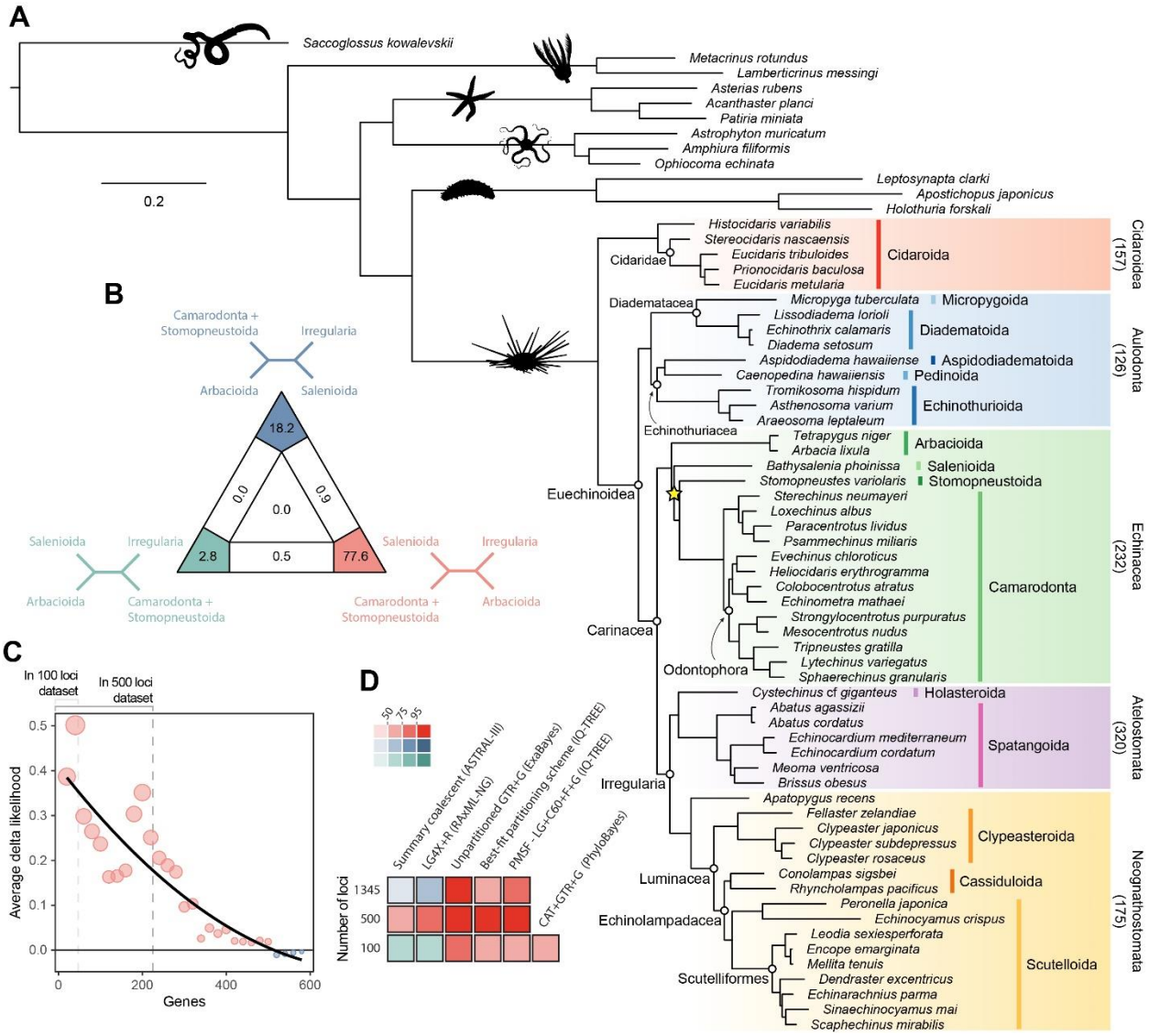


Fig. 2: Phylogenetic relationships among major clades of Echinoidea. **A.** Favored topology. With the exception of a single contentious node within Echinacea (marked with a star), all methods supported the same pattern of relationships. Numbers below major clades correspond to the current numbers of 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 clade of Salenioida + (Camarodonta + Stomopneustoida), as depicted in **A**, is predominantly concentrated among the most

156 phylogenetically useful loci. This signal is attenuated in larger datasets that contain less reliable genes,
157 eventually favoring an alternative resolution. Only the 584 loci containing data for the three main
158 lineages of Echinacea were considered. The line corresponds to a second-degree polynomial regression.
159 **D.** Resolution and bootstrap scores (color coded) of the topology within Echinacea found using datasets
160 of different sizes and alternative methods of inference.

161

162 Salenioida is another major lineage sampled here for the first time, and whose exact position
163 among regular echinoids proved difficult to resolve. While some methods supported salenioids as the
164 sister group to a clade of camarodonts, stomopneustoids, and arbacioids (a topology previously
165 supported by morphology [19]), others recovered a closer relationship of salenioids to Camarodonta +
166 Stomopneustoida, with arbacioids sister to them all (as shown in Fig. 2A). As revealed using likelihood
167 mapping, these results do not stem from a lack of phylogenetic signal, but rather from the presence of
168 strong and conflicting evidence in the dataset regarding the position of salenioids (Fig. 2B). However, a
169 careful dissection of these signals shows that loci with high phylogenetic usefulness (as defined by [3,
170 37]; see Methods) favor the topology shown in Fig. 2A, with the morphological hypothesis becoming
171 dominant only after incorporating less reliable loci (Fig .2C). In line with these results, moderate levels of
172 gene subsampling (down to 500 loci) targeting the most phylogenetically useful loci unambiguously
173 support the placement of arbacioids as sister to the remaining taxa, regardless of the chosen method of
174 inference (Fig. 2D). More extreme subsampling (down to 100 loci) again results in disagreement among
175 methods. This possibly stems from the increasing effect of stochastic errors in smaller datasets, as less
176 than half of the sampled loci in these reduced datasets contain data for all branches of this quartet (see
177 Fig. 2C). This result shows the importance of ensuring that datasets (especially subsampled ones) retain
178 appropriate levels of occupancy for clades bracketing contentious nodes [38]. Despite these
179 disagreements, several lines of evidence favor the topology shown in Fig. 2A, including the results of

180 likelihood mapping, and the increased support for this resolution among the most phylogenetically
181 useful loci and when using more complex methods of reconstruction, such as partitioned and site-
182 heterogeneous models, which favor this topology regardless of dataset size (Fig. 2D).

183 While alternative methods of inference had minor effects on phylogenetic relationships, they
184 did impact the reconstruction of branch lengths (Fig. S1). Site-heterogeneous models (such as
185 CAT+GTR+G) returned longer branch lengths overall, but also uncovered a larger degree of molecular
186 change among echinoderm classes. Branches connecting these clades were stretched to a much larger
187 extent than those within the ingroup (Fig. S1), which might affect the inference of node dates. We
188 tested this hypothesis by exploring the sensitivity of divergence times to the use of alternative models of
189 molecular evolution (site-homogeneous vs. site-heterogeneous), as well as different clock priors
190 (autocorrelated vs. uncorrelated), and gene sampling strategies (using five different approaches; see
191 Methods). All combinations of these factors were explored, resulting in 20 different time-calibration
192 settings that were run under Bayesian approaches. Inferred divergence times were significantly affected
193 by all of these factors (MANOVA; all $P < 2.2 \times 10^{-16}$). While nodes connecting some outgroup taxa were
194 among those most sensitive to these decisions, large effects were also seen among nodes relating to the
195 origin and diversification of the echinoid clades Cidaroidea and Aulodonta (Fig. 3A). All of these nodes
196 varied in age by more than 60 Ma—and up to 115 Ma—among the consensus topologies of different
197 analyses (Fig. 3A).

198

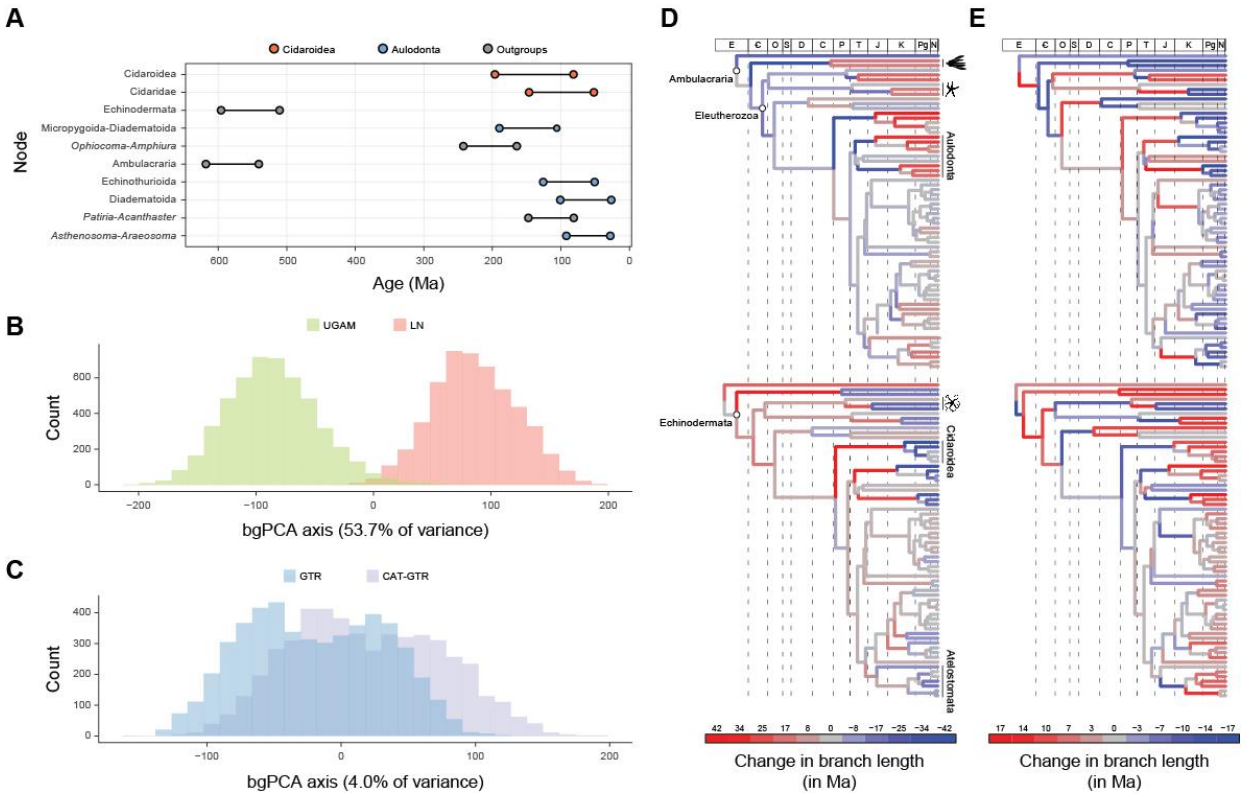


Fig. 3: Sensitivity of divergence time estimation to methodological decisions. **A.** The ten most sensitive ages are found within Cidaroidea, Aulodonta, and among outgroup nodes. For each, the range shown spans the interval between the minimum and maximum ages found among the consensus topologies of all 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**).

211 In order to isolate and visualize the impact of these factors on divergence time estimation, we
212 extracted the recovered node ages from the posterior distribution of time-calibrated topologies under
213 different settings. This allowed chronograms to be represented in a multidimensional space of node
214 dates, which we term a chronospace, given its similarities to treespaces commonly used to explore
215 topological differences [39]. Each observation was also classified as obtained under a specific clock prior,
216 model of molecular evolution, and gene sampling strategy, and the major effects of each of these
217 choices were extracted with the use of between-group principal component analyses (bgPCAs). This
218 revealed that the single dimension of chronospace maximizing the distinctiveness of trees obtained
219 under different clock models explained 53.7% of the total variance in node ages across all analyses (Figs.
220 3B-E and S2). In contrast, the choice of different loci or models of molecular evolution showed a much
221 lesser effect, explaining only 10.5% and 4.0% of the total variance, respectively. Even though all of these
222 decisions affected a similar set of sensitive nodes (those mentioned above, as well as some relationships
223 within Atelostomata), the choice of clock model modified the ages of 17 of these by more than 20 Ma
224 (Fig. S3). This degree of change was induced on only 5 nodes by selecting alternative loci (Fig. S4), and
225 was not induced on any node by enforcing different models of evolution (Fig. S5).

226 Even when the age of crown Echinodermata was constrained to postdate the appearance of
227 stereom (the characteristic skeletal microstructure of echinoderms) in the Early Cambrian [40, 41], only
228 analyses using the most clock-like loci recovered ages concordant with this (i.e., median ages younger
229 than the calibration enforced; Fig. S6). Instead, most consensus trees favored markedly older ages for
230 the clade, in some cases even predating the origin of the Ediacaran biota [42] (Fig. S6). Despite the
231 relative sensitivity of many of the earliest nodes to methodological choices (Fig. 3A), the split between
232 Crinoidea and all other echinoderms (Eleutherozoa) is always inferred to have predated the end of the
233 Cambrian (youngest median age = 492.2 Ma), and the divergence among the other major lineages
234 (classes) of extant echinoderms are constrained to have happened between the Late Cambrian and

Middle Ordovician (Fig. S6). Our results also recover an early origin of crown group Holothuroidea (sea cucumbers; range of median ages = 351.6–383.2 Ma), well before the crown groups of other extant echinoderm classes. These dates markedly postdate the first records of holothuroid calcareous rings in the fossil record [43, 44], and imply that this trait does not define the holothuroid crown group but instead evolved from an echinoid-like jaw-apparatus along its stem [45]. The other noteworthy disagreement between these results and those of previous studies [46] 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).

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 (Fig. 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 (Fig. S7). 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 (Figs. 4 and S6). 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 (Fig. 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 (Fig. S7).

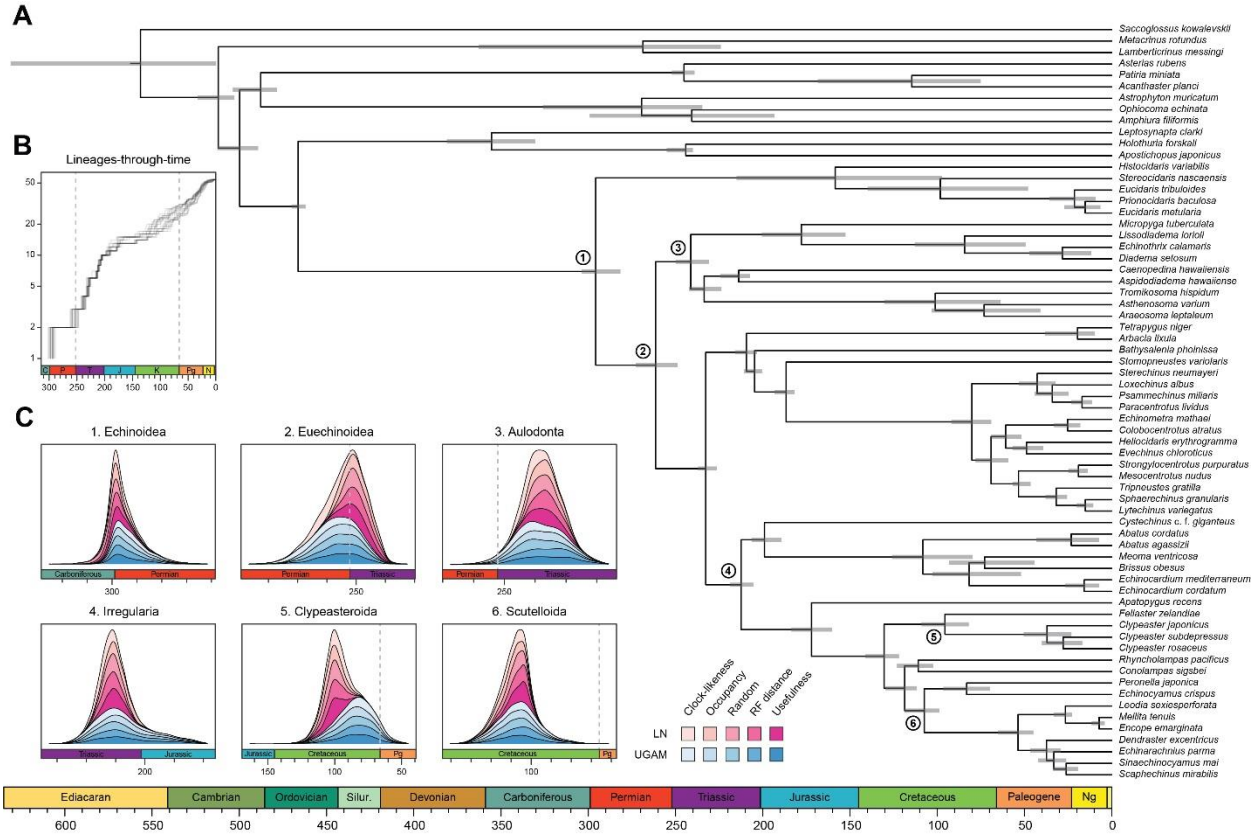


Fig. 4: Divergence times among major clades of Echinoidea and other echinoderms. **A.** Consensus chronogram of the two runs using clock-like genes under a CAT+GTR+G model of molecular evolution and a autocorrelated log-normal (LN) clock prior. Node ages correspond to median values, and bars show the 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 an individual consensus topology from among the time-calibrated runs performed. **C.** Posterior distributions of the ages of selected nodes (identified in **A**). The effects of models of molecular evolution are not shown, as they represent the least important factor (see Fig. 3); distributions using site-homogeneous and heterogeneous models are combined for every combination of targeted loci and clock prior.

Discussion

(a) The echinoid Tree of Life

In agreement with previous phylogenomic studies [3, 4], echinoid diversity can be subdivided into five major clades (Fig. 2A). Cidaroids form the sister group to all other crown group echinoids (Euechinoidea). Some aspects of the relationships among sampled cidaroids are consistent with previous molecular [47] and morphological studies [19], including an initial split between *Histocidaris* and the remaining taxa, representing the two main branches of extant cidaroids [5, 35]. Others, such as the nested position of *Prionocidaris baculosa* within the genus *Eucidaris,* not only implies paraphyly of this genus but also suggests the need for a taxonomic reorganization of the family Cidaridae. Within euechinoids, the monophyly of Aulodonta is supported for the first time with sampling of all of its major groups. The subdivision of these into a clade that includes diadematoids plus micropygoids (which we propose should retain the name Diadematacea), sister to a clade including echinothurioids and pedinoids (Echinothuriacea *sensu* [4]) is strongly reminiscent of some early classifications (e.g., [48]). Our expanded phylogenomic sampling also confirms an aulodont affinity for aspidodiadematoids [3, 35] and places them within Echinothuriacea as the sister group to Pedinoida.

The remaining diversity of echinoids, which forms the clade Carinacea (Fig. 2), is subdivided into Irregularia and their sister clade among regulars, for which we amend the name Echinacea to include Salenioida. Given the striking morphological gap separating regular and irregular echinoids, the origin of Irregularia has been shrouded in mystery [19, 23, 48]. Our complete sampling of major regular lineages determines Echinacea *sensu stricto* to be the sister clade to irregular echinoids. A monophyletic Echinacea was also supported in a recent total-evidence analysis [3] , but the incomplete molecular sampling of that study resulted in a slightly different topology that placed salenioids as the sister group

294 to the remaining lineages. However, an overall lack of morphological synapomorphies uniting these
295 clades had previously been acknowledged [19]. While the relationships within Echinacea proved to be
296 difficult to resolve even with thousands of loci, multiple lines of evidence lead us to prefer a topology in
297 which salenioids form a clade with camarodonts + stomopneustoids, with arbacioids sister to all of these
298 (Fig. 2).

299 As has been already established [3, 4, 14, 19, 30], the lineages of irregular echinoids here
300 sampled are subdivided into Atelostomata (heart urchins and allies) and Neognathostomata (sand
301 dollars, sea biscuits, and "cassiduloids"). Despite the former being the most diverse of the five main
302 clades of echinoids (Fig. 2), its representation in phylogenomic studies remains low, and its internal
303 phylogeny poorly constrained [35]. On the contrary, recent molecular studies have greatly improved our
304 understanding of the relationships among neognathosomates [3, 4, 36], revealing an evolutionary
305 history that dramatically departs from previous conceptions. Even when scutelloids and clypeasteroids
306 were never recovered as reciprocal sister lineages by molecular phylogenies (e.g., [13, 14, 30]), this
307 result was not fully accepted until phylogenomic data confidently placed echinolampadids as the sister
308 lineage to sand dollars [4]. At the same time, this result rendered the position of the remaining
309 "cassiduloids", a taxonomic wastebasket with an already complicated history of classification [19, 27, 28,
310 49], entirely uncertain. An attempt to constrain the position of these using a total-evidence approach [3]
311 subdivided the "cassiduloids" into three unrelated clades: Nucleolitoida, composed of extinct lineages
312 and placed outside the node defined by Scutelloida + Clypeasteroida, and two other clades nested
313 within it (see Fig. 1). Extant "cassiduloids" were recovered as members of one of the latter clades,
314 representing the monophyletic sister group to sand dollars. Here, we show that *Apatopygus recens* does
315 not belong within this clade but is instead the sister group to all other extant neognathostomates. Given
316 this phylogenetic position, as well as the morphological similarities between *Apatopygus* and the
317 entirely extinct nucleolitids [19, 28, 49-51], it is likely that the three extant species of apatopygids

318 represent the last surviving remnants of Nucleolitoida, a clade of otherwise predominantly Mesozoic
319 neognathostomates [3]. Because of the renewed importance in recognizing this topology, we propose
320 the name Luminacea for the clade uniting all extant neognathostomates with the exclusion of
321 *Apatopygus* (Fig. 2). This nomenclature refers to the dynamic evolutionary history of the Aristotle's
322 lantern (i.e., the echinoid jaw-apparatus) within the clade (present in the adults of both clypeasteroids
323 and scutelloids, but found only in the juveniles of Cassiduloida *sensu stricto*), the inclusion of the so-
324 called lamp urchins (echinolampadids) within the clade, and the illumination provided by this hitherto
325 unexpected topology. The previous misplacement of *Apatopygus* within this clade ([3]; see Fig. 1G) is
326 likely a consequence of tip-dating preferring more stratigraphically congruent topologies [52], an effect
327 that can incorrectly resolve taxa on long terminal branches [53]. Given the generally useful phylogenetic
328 signal of stratigraphic information [54], this inaccuracy further highlights the unusual evolutionary
329 history of living apatopygids.

330

331 **(b) Echinoid diversification**

332 Calibrating phylogenies to absolute time is crucial to understanding evolutionary history, as the
333 resulting chronograms provide a major avenue for testing hypotheses of diversification. However, the
334 accuracy and precision of the inferred divergence times hinge upon many methodological choices that
335 are often difficult or time-consuming to justify (e.g., calibration strategies, prior distributions on node
336 ages, clock models, etc.) [55-60]. Understanding the individual impacts of each of these sources of
337 uncertainty can be difficult. Sensitivity of inferred ages is commonly explored by running analyses under
338 different settings and summarizing results in either tables or by stacking chronograms in order to
339 visualize the relative position of nodes (see for example [56, 57]). This approach discards a lot of
340 information contained in Bayesian posterior distributions of time-calibrated trees. For example, patterns

341 of correlation among branch lengths (i.e., ways in which changes in a given branch length are
342 accommodated by other regions of the tree) are lost, and alternative ways in which node ages are
343 attained (e.g., which of all descendant branches are expanding as a given node becomes older) are not
344 usually explored, even though they can potentially represent very different evolutionary scenarios.

345 Here we explore the sensitivity of node ages to different ways of modelling heterogeneities
346 among lineages in the rates and patterns of molecular evolution, as well as alternative criteria to sample
347 loci. To do so, we introduce an approach to visualize the distribution of chronograms in a
348 multidimensional space of node ages (a chronospace), and measure the overall effect of these decisions
349 on inferred dates using multivariate statistical methods. Our results reveal that divergence times
350 obtained under site-homogeneous and site-heterogeneous models (such as CAT+GTR+G) are broadly
351 comparable. This happens despite these types of models estimating higher levels of sequence
352 divergence and stretching branches in a non-isometric manner (Fig. S1). A similar result was recently
353 found when comparing site-homogeneous models with different numbers of parameters [61],
354 suggesting that relaxed clocks adjust branch rates in a manner that buffers the effects introduced by
355 using more complex (and computationally-intensive) models of evolution. Similarly, the choice between
356 different loci also has a small effect on inferred ages, with little evidence of a systematic difference
357 between the divergence times supported by randomly-chosen loci and those found using targeted
358 sampling criteria, such as selecting genes for their phylogenetic signal, usefulness, occupancy, or clock-
359 likeness. A meaningful effect was restricted to a few ancient nodes (e.g., Echinodermata), for which
360 clock-like genes suggested younger ages that are more consistent with fossil evidence. While this
361 validates the use of clock-like genes for inferring deep histories of diversification, the choice of loci had
362 no meaningful effect on younger ages. Finally, the choice between alternative clock models induced
363 differences in ages that were five to ten times stronger than those of other factors, emphasizing the

importance of either validating their choice (e.g., [62]) or—as done here—focusing on results that are robust to them.

The origin and early diversification of crown group Echinoidea have always been considered to have been determined (or at least strongly affected) by the P-T mass extinction [11, 12, 16, 20]. However, estimating the number of crown group members surviving the most severe biodiversity crisis in the Phanerozoic [63] has been hampered by both paleontological and phylogenetic uncertainties [3, 10, 13-15, 18]. Our results establish with confidence that multiple crown group lineages survived and crossed this boundary, finding for the first time a null posterior probability of the clade originating after the extinction event. While the survival of three crown group lineages is slightly favored (Fig. S8), discerning between alternative scenarios is still precluded by uncertainties in dating these early divergences. Echinoid diversification during the Triassic was relatively fast (Fig. 4B) and involved rapid divergences among its major clades. Even many lineages presently classified at the ordinal level trace their origins to this initial pulse of diversification following the P-T mass extinction.

The late Paleozoic and Triassic origins inferred for the crown group and many euechinoid orders prompts a re-evaluation of fossils from this interval of time. Incompletely known fossil taxa such as the Pennsylvanian *?Eotiaris meurevillensis*, with an overall morphology akin to that of crown group echinoids, has a stratigraphic range consistent with our inferred date for the origin of the echinoid crown group [20]. Additionally, the Triassic fossil record of echinoids has been considered to be dominated by stem group cidaroids, with the first euechinoids not known until the Late Triassic [17, 64]. However, the Triassic origins of many euechinoid lineages supported by our analyses necessitates that potential euechinoid affinities should be re-considered for this diversity of Triassic fossils. This is especially the case for the serpianotiarids and triadocidarids, abundant Triassic families variously interpreted as cidaroids, euechinoids, or even stem echinoids [3, 17, 64, 65]. A reinterpretation of any of these as euechinoids would suggest that the long-implied gap in the euechinoid record [16, 18] is caused

388 by our inability to correctly place these key fossils, as opposed to an incompleteness of the fossil record
389 itself.

390 While our phylogenomic approach is the first to resolve the position of all major cassiduloid
391 lineages, the inferred ages for many nodes within Neognathostomata remain in strong disagreement
392 with the fossil record. No Mesozoic fossil can be unambiguously assigned to either sand dollars or sea
393 biscuits, a surprising situation given the good fossilization potential and highly distinctive morphology of
394 these clades [19, 21, 25, 66]. While molecular support for a sister group relationship between scutelloids
395 and echinolampadids already implied this clade (Echinolampadacea) must have split from clypeasteroids
396 by the Late Cretaceous [3, 4, 14, 19], this still left open the possibility that the crown groups of sand
397 dollars and sea biscuits radiated in the Cenozoic. Under this scenario, the Mesozoic history of these
398 groups could have been comprised of forms lacking their distinctive morphological features,
399 complicating their correct identification. This hypothesis is here rejected, with the data unambiguously
400 supporting the radiation of both crown groups preceding the K-Pg mass extinction (Fig. 4C). While it
401 remains possible that these results are incorrect even after such a thorough exploration of the time-
402 calibration toolkit (see for example [60, 67]), these findings call for a critical reassessment of the
403 Cretaceous fossil record, and a better understanding of the timing and pattern of morphological
404 evolution among fossil and extant neognathostomates. For example, isolated teeth with an overall
405 resemblance to those of modern sand dollars and sea biscuits have been found in Lower Cretaceous
406 deposits [68], raising the possibility that other overlooked and disarticulated remains might close the
407 gap between rocks and clocks.

408

409 **(c) Conclusions**

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

419 Although divergence time estimation is known to be sensitive to many methodological
420 decisions, systematically quantifying the relative impact of these on inferred ages has rarely been done.
421 Here we propose an approach based on chronospaces (implemented in R [69] and available at
422 https://github.com/mongiardino/chronospace) that can help visualize key effects and determine the
423 sensitivity of node dates to different assumptions. Our results shed new light on the early evolutionary
424 history of crown group echinoids and its relationship with the P-T mass extinction event, a point in time
425 where the fossil record provides ambiguous answers. They also establish with confidence a Cretaceous
426 origin for the sand dollars and sea biscuits, preceding their first appearance in the fossil record by at
427 least 40 to 50 Ma, respectively (and potentially up to 65 Ma). These clades, therefore, join several well-
428 established cases of discrepancies between the fossil record and molecular clocks, such as those
429 underlying the origins of placental mammals [70] and flowering plants [71].

430

431 **Methods**

432 **(a) Sampling, bioinformatics, and matrix construction**

433 This study builds upon previous phylogenomic matrices [3, 4], the last of which was augmented
434 through the addition of eight published datasets (mostly expanding outgroup sampling), as well as 17
435 novel echinoid datasets (15 transcriptomes and two draft genomes). For all novel datasets, tissue
436 sampling, DNA/RNA extraction, library preparation, and sequencing varied by specimen, and are
437 detailed in the Supplementary Information. Raw reads have been deposited in NCBI under Bioproject
438 accession numbers XXX and XXX. Final taxon sampling included 12 outgroups and 50 echinoids
439 (accession numbers and sampling details can be found in Table S1).

440 Raw reads for all transcriptomic datasets were trimmed or excluded using quality scores with
441 Trimmomatic v. 0.36 [72] under default parameters. Further sanitation steps were performed using the
442 Agalma 2.0 phylogenomic workflow [73], and datasets were assembled *de novo* with Trinity 2.5.1 [74].
443 For genomic shotgun sequences, adapters were removed with BBDuk
444 (https://sourceforge.net/projects/bbmap/), and UrQt v. 1.0.18 [75] was used to filter short reads (size <
445 50) and trim low-quality ends (score < 28). Datasets were then assembled using MEGAHIT v. 1.1.2 [76].
446 Draft genomes were masked using RepeatMasker v. 4.1.0 [77, 78], before obtaining gene predictions
447 with AUGUSTUS 3.2.3 [79]. A custom set of universal single-copy orthologs (USCOs) obtained from the
448 latest *Strongylocentrotus purpuratus* genome assembly (Spur v. 5.0) was employed as the training
449 dataset. Settings and further details of these analyses can be found in the Supplementary Information.

450 Multiplexed transcriptomes were sanitized from cross-contaminants using CroCo v. 1.1 [80], and
451 likely non-metazoan contaminants were removed using alien_index v. 3.0 [81] (removing sequences
452 with AI values > 45). Datasets were imported back into Agalma, which automated orthology inference
453 (as described in [73, 82]), gene alignment with MACSE [83], and trimming with GBLOCKS [84]. The amino
454 acid supermatrix was reduced using a 70% occupancy threshold, producing a final dataset of 1,346 loci

455 (327,695 sites). As a final sanitation step, gene trees were obtained using ParGenes v. 1.0.1 [85], which
456 performed model selection (minimizing the Bayesian Information Criterion) and inference using 100
457 bootstrap (BS) replicates. Trees were then used to remove outlier sequences with TreeShrink v. 1.3.1
458 [86]. We specified a reduced tolerance for false positives and limited removal to at most three terminals
459 which had to increase tree diameter by at least 25% (-q 0.01 -k 3 -b 25). Statistics for the supermatrix
460 and all assemblies can be found in Table S2.

461

462 **(b) Phylogenetic inference**

463 Inference was performed under multiple probabilistic and coalescent-aware methods, known to
464 differ in their susceptibility to model violations. Coalescent-based inference was performed using the
465 summary method ASTRAL-III [87], estimating support as local posterior probabilities [88]. Among
466 concatenation approaches, we used Bayesian inference under an unpartitioned GTR+G model in
467 ExaBayes 1.5 [89]. Two chains were run for 2.5 million generations, samples were drawn every one
468 hundred, and the initial 25% was discarded as burn-in. We also explored maximum likelihood inference
469 with partitioned and unpartitioned models. For the former, the fast-relaxed clustering algorithm was
470 used to find the best-fitting model among the top 10% using IQ-TREE 1.6.12 [90, 91] (-m MFP+MERGE -
471 rclusterf 10 -rcluster-max 3000), and support was evaluated with 1,000 ultrafast BS replicates [92]. For
472 the latter, we used the LG4X+R model in RAxML-NG v. 0.5.1 [93] and evaluated support with 200
473 replicates of BS. Finally, we also implemented the site-heterogeneous LG+C60+F+G mixture model using
474 the posterior mean site frequency (PMSF) approach to provide a fast approximation of the full profile
475 mixture model [94], allowing the use of 100 BS replicates to estimate support. Given some degree of
476 topological conflict between the results of the other methods (see below), multiple guide trees were
477 used to estimate site frequency profiles, but the resulting phylogenies were identical.

478 Given conflicts between methods in the resolution of one particular node (involving the
479 relationships among Arbacioida, Salenioida, and the clade of Stomopneustoida + Camarodonta), all
480 methods were repeated after reducing the matrix to 500 and 100 loci selected for their phylogenetic
481 usefulness using the approach described in [3, 37] and implemented in the *genesortR* script
482 (https://github.com/mongiardino/genesortR). This approach relies on seven gene properties routinely
483 used for phylogenomic subsampling, including multiple proxies for phylogenetic signal—such as the
484 average BS and Robinson-Foulds (RF) similarity to a target topology—as well as several potential sources
485 of systematic bias (e.g., rate and compositional heterogeneities). Outgroups were removed before
486 calculating these metrics. RF similarity was measured to a species tree that had the conflicting
487 relationship collapsed so as not to bias gene selection in favor of any resolution. A principal component
488 analysis (PCA) of this dataset resulted in a dimension (PC 2, 17.6% of variance) along which phylogenetic
489 signal increased while sources of bias decreased (Fig. S9), and which was used for loci selection. For the
490 smallest dataset, we also performed inference under the site-heterogeneous CAT+GTR+G model using
491 PhyloBayes-MPI [95]. Three runs were continued for > 10,000 generations, sampling every two
492 generations and discarding the initial 25%. Convergence was confirmed given a maximum bipartition
493 discrepancy of 0.067 and effective sample sizes for all parameters > 150.

494 Two other approaches were implemented in order to assist in resolving the contentious node.
495 First, we implemented a likelihood-mapping analysis [96] in IQ-TREE to visualize the phylogenetic signal
496 for alternative resolutions of the quartet involving these three lineages (Arbacioida, Salenioida, and
497 Stomopneustoida + Camarodonta) and their sister clade (Irregularia; other taxa were excluded). Second,
498 we estimated the log-likelihood scores of each site in RAxML (using best-fitting models) for the two most
499 strongly supported resolutions found through likelihood mapping. These were used to calculate gene-
500 wise differences in scores, or δ values [97]. In order to search for discernable trends in the signal for
501 alternative topologies, genes were ordered based on their phylogenetic usefulness (see above) and the

502 mean per-locus δ values of datasets composed of multiples of 20 loci (i.e., the most useful 20, 40, etc.)
503 were calculated.

504

505 **(c) Time calibration**

506 Node dating was performed using relaxed molecular clocks in PhyloBayes v4.1 using a fixed
507 topology and a novel set of 22 fossil calibrations corresponding to nodes from our newly inferred
508 phylogeny (listed in the Supplementary Information). Depending on the node, we enforced both
509 minimum and maximum bounds, or either one of these. A birth-death prior was used for divergence
510 times, which allowed for the implementation of soft bounds [98], leaving 5% prior probability of
511 divergences falling outside of the specified interval. We explored the sensitivity of divergence time
512 estimates to gene selection, model of molecular evolution, and clock prior. One hundred loci were
513 sampled from the full supermatrix according to four targeted sampling schemes: usefulness (calculated
514 as explained above, except incorporating all echinoderm terminals), phylogenetic signal (i.e., smallest RF
515 distance to species tree), clock-likeness (i.e., smallest variance of root-to-tip distances), and level of
516 occupancy. For clock-likeness, we only considered loci that lay within one standard deviation of the
517 mean rate (estimated by dividing total tree length by the number of terminals [99]), as this method is
518 otherwise prone to selecting largely uninformative loci (Fig. S10; [37]). A fifth sample of randomly
519 chosen loci was also evaluated. These five datasets were run under two models of molecular evolution,
520 the site-homogeneous GTR+G and the site-heterogeneous CAT+GTR+G, and both uncorrelated gamma
521 (UGAM) and autocorrelated log-normal (LN) clocks were implemented. The combination of these
522 settings (loci sampled, model of evolution, and clock prior) resulted in 20 analyses. For each, two runs
523 were continued for 20,000 generations, after which the initial 25% was discarded and the chains thinned
524 to every two generations (see log-likelihood trace plots in Fig. S11). To explore the sensitivity of

525 divergence times to these methodological decisions, 500 random chronograms were sampled from each
526 analysis (250 from each run), and their node dates were subjected to multivariate analyses of variance
527 (MANOVA) and between-group PCA (bgPCA) using package *Morpho* [100] in the R statistical
528 environment [69]. bgPCA involves the use of PCA on the covariance matrix of group means (e.g., the
529 mean ages of all nodes obtained under UGAM or LN clocks), followed by the projection of original
530 samples onto the obtained bgPC axes. The result is a multidimensional representation of divergence
531 times—a chronospace—rotated so as to capture the distinctiveness of observations obtained under
532 different settings. Separate bgPCAs were performed for loci sampling strategy, clock prior and model of
533 molecular evolution, and the proportion of total variance explained by bgPC axes was interpreted as an
534 estimate of the relative impact of these choices on inferred times of divergence. Finally, lineage-
535 through-time plots were generated using *ape* [101].

536

537 **Acknowledgments**

538 We are grateful to the crew of R/Vs Antéa, Atlantis, Falkor and Mirai, HOV Alvin, and the crew
539 and PIs of BIOMAGLO deep-sea cruises (Laure Corbari, Karine Olu-Le Roy, Sarah Samadi) for assistance in
540 specimen collection. We would also like to thank Libby Liggins, Wilma Blom, Owen Anderson, Francisco
541 Solís-Marín, Carlos A. Conejeros-Vargas, and Jih-Pai Lin for the collection and shipment of specimens.
542 We gratefully acknowledge assistance from Thomas Saucède (Univ. Burgundy), Marc Eléaume (MNHN),
543 and Charlotte Seid (Scripps) in accessing, cataloging and processing museum specimens. Tim Ravasi
544 provided resources and collection facilities to GWR via King Abdullah University of Science and
545 Technology (KAUST). Institutional support was provided by the Central Research Laboratories and the
546 Department of Geology and Palaeontology at the Natural History Museum Vienna, Austria. This work
547 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.

Competing interests

We declare no competing interests.

References

1. Donoghue PC, Benton MJ. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. Trends Ecol Evol. 2007;22(8):424-431.
2. Smith AB, Peterson KJ. Dating the time of origin of major clades: molecular clocks and the fossil record. Annu Rev Earth Planet Sci. 2002;30(1):65-88.
3. Mongiardino Koch N, Thompson JR. A total-evidence dated phylogeny of Echinoidea combining phylogenomic and paleontological data. Syst Biol. 2021;70(3):421-439.
4. Mongiardino Koch N, Coppard SE, Lessios HA, Briggs DE, Mooi R, Rouse GW. A phylogenomic resolution of the sea urchin tree of life. BMC Evol Biol. 2018;18(1):189.
5. Kroh A, Mooi R. World Echinoidea Database; 2021 [cited 2021 July 13]. Available from: http://www.marinespecies.org/echinoidea.
6. Schultz HA. Echinoidea: with pentameral symmetry: Walter de Gruyter GmbH & Co KG; 2015.

7. Edmunds PJ, Carpenter RC. Recovery of *Diadema antillarum* reduces macroalgal cover and increases abundance of juvenile corals on a Caribbean reef. Proc Natl Acad Sci. 2001;98(9):5067-5071.

8. Harrold C, Pearse JS. The ecological role of echinoderms in kelp forests. In: Jangoux M, Lawrence JM, editors. Echinoderm Studies 2. Rotterdam: AA Balkema; 1987. pp. 137–233.

9. Zonneveld JP, Furlong CM, Sanders SC. Triassic echinoids (Echinodermata) from the Aksala Formation, north Lake Laberge, Yukon Territory, Canada. Pap Palaeontol. 2016;2(1):87-100.

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

11. Kier PM. Triassic echinoids. Smithson Contrib Paleobiol. 1977;30:1-88.

12. Twitchett RJ, Oji T. Early Triassic recovery of echinoderms. C R Palevol. 2005;4(6-7):531-542.

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

14. Smith AB, Pisani D, Mackenzie-Dodds JA, Stockley B, Webster BL, Littlewood DTJ. Testing the molecular clock: molecular and paleontological estimates of divergence times in the Echinoidea (Echinodermata). Mol Biol Evol. 2006;23(10):1832-1851.

15. Smith A, Hollingworth N. Tooth structure and phylogeny of the Upper Permian echinoid *Miocidaris keyserlingi*. Proc Yorkshire Geol Soci. 1990;48(1):47-60.

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

591 17. Smith AB. Intrinsic versus extrinsic biases in the fossil record: contrasting the fossil record of
592 echinoids in the Triassic and early Jurassic using sampling data, phylogenetic analysis, and molecular
593 clocks. Paleobiology. 2007;33(2):310-323.
594 18. Thompson JR, Hu S-x, Zhang Q-Y, Petsios E, Cotton LJ, Huang J-Y, et al. A new stem group
595 echinoid from the Triassic of China leads to a revised macroevolutionary history of echinoids during the
596 end-Permian mass extinction. R Soc Open Sci. 2018;5(1):171548.
597 19. Kroh A, Smith AB. The phylogeny and classification of post-Palaeozoic echinoids. J Syst
598 Palaeontol. 2010;8(2):147-212.
599 20. Thompson JR, Mirantsev GV, Petsios E, Bottjer DJ. Phylogenetic analysis of the Archaeocidaridae
600 and Palaeozoic Miocidaridae (Echinodermata, Echinoidea) and the origin of crown group echinoids. Pap
601 Palaeontol. 2020;6(2):217-249.
602 21. Kier PM. Rapid evolution in echinoids. Palaeontology. 1982;25(1):1-9.
603 22. Barras CG. Morphological innovation associated with the expansion of atelostomate irregular
604 echinoids into fine-grained sediments during the Jurassic. Palaeogeogr Palaeoclimatol Palaeoecol.
605 2008;263(1-2):44-57.
606 23. Saucède T, Mooi R, David B. Phylogeny and origin of Jurassic irregular echinoids (Echinodermata:
607 Echinoidea). Geol Mag. 2007;144(2):333-359.
608 24. Hopkins MJ, Smith AB. Dynamic evolutionary change in post-Paleozoic echinoids and the
609 importance of scale when interpreting changes in rates of evolution. Proc Natl Acad Sci.
610 2015;112(12):3758-3763.
611 25. Mooi R. Paedomorphosis, Aristotle's lantern, and the origin of the sand dollars (Echinodermata:
612 Clypeasteroida). Paleobiology. 1990;16(1):25-48.
613 26. Nebelsick JH. Ecology of clypeasteroids. In: Lawrence JM, editor. Sea Urchins: Biology and
614 Ecology, 4th edition. Cambridge: Academic Press; 2020. pp. 315-331.

615 27. Smith AB. Probing the cassiduloid origins of clypeasteroid echinoids using stratigraphically
616 restricted parsimony analysis. Paleobiology. 2001;27(2):392-404.
617 28. Souto C, Mooi R, Martins L, Menegola C, Marshall CR. Homoplasy and extinction: the phylogeny
618 of cassidulid echinoids (Echinodermata). Zool J Linn Soc. 2019;187:622-660.
619 29. Kier PM. Revision of the cassiduloid echinoids. Smithson Misc Collect. 1962;144:1-262.
620 30. Littlewood D, Smith A. A combined morphological and molecular phylogeny for sea urchins
621 (Echinoidea: Echinodermata). Philos Trans R Soc. 1995;347(1320):213-234.
622 31. Mongiardino Koch N. Exploring adaptive landscapes across deep time: A case study using
623 echinoid body size. Evolution. 2021;75:1567-1581.
624 32. Mooi R. Progenetic miniaturization in the sand dollar *Sinaechinocyamus*: implications for
625 clypeasteroid phylogeny. In: De Ridder C, Dubois P, Lahaye M, Jangoux M, editors. Echinoderm
626 Research. Rotterdam: AA Balkema; 1990. pp. 137-143.
627 33. Dunn CW, Zapata F, Munro C, Siebert S, Hejnol A. Pairwise comparisons across species are
628 problematic when analyzing functional genomic data. Proc Natl Acad Sci. 2018;115(3):409-417.
629 34. Smith SD, Pennell MW, Dunn CW, Edwards SV. Phylogenetics is the new genetics (for most of
630 biodiversity). Trends Ecol Evol. 2020;35(5):415-425.
631 35. Kroh A. Phylogeny and classification of echinoids. In: Lawrence J, editor. Sea Urchins: Biology
632 and Ecology, 4th edition. Cambridge: Academic Press; 2020. p. 1-17.
633 36. Lin J-P, Tsai M-H, Kroh A, Trautman A, Machado DJ, Chang L-Y, et al. The first complete
634 mitochondrial genome of the sand dollar *Sinaechinocyamus mai* (Echinoidea: Clypeasteroida).
635 Genomics. 2020;112(2):1686-1693.
636 37. Mongiardino Koch N. Phylogenomic subsampling and the search for phylogenetically reliable
637 loci. Mol Biol Evol. 2021;msab151. doi: 10.1093/molbev/msab151.

638 38. Dell'Ampio E, Meusemann K, Szucsich NU, Peters RS, Meyer B, Borner J, et al. Decisive data sets
639 in phylogenomics: lessons from studies on the phylogenetic relationships of primarily wingless insects.
640 Mol Biol Evol. 2014;31(1):239-249.
641 39. Hillis DM, Heath TA, John KS. Analysis and visualization of tree space. Syst Biol. 2005;54(3):471-
642 482.
643 40. Zamora S, Lefebvre B, Álvaro JJ, Clausen S, Elicki O, Fatka O, et al. Cambrian echinoderm
644 diversity and palaeobiogeography. Geol Soc Lond Mem. 2013;38(1):157-171.
645 41. Bottjer DJ, Davidson EH, Peterson KJ, Cameron RA. Paleogenomics of echinoderms. Science.
646 2006;314(5801):956-960.
647 42. Pu JP, Bowring SA, Ramezani J, Myrow P, Raub TD, Landing E, et al. Dodging snowballs:
648 Geochronology of the Gaskiers glaciation and the first appearance of the Ediacaran biota. Geology.
649 2016;44(11):955-958.
650 43. Reich M. Different pathways in early evolution of the holothurian calcareous ring. In: Zamora S,
651 Rábano I, editors. Progress in Echinoderm Palaeobiology. Madrid: Instituto Geológico y Minero de
652 España; 2015, pp. 137-145.
653 44. Miller AK, Kerr AM, Paulay G, Reich M, Wilson NG, Carvajal JI, et al. Molecular phylogeny of
654 extant Holothuroidea (Echinodermata). Mol Phylogenet Evol. 2017;111:110-131.
655 45. Rahman IA, Thompson JR, Briggs DE, Siveter DJ, Siveter DJ, Sutton MD. A new ophiocistioid with
656 soft-tissue preservation from the Silurian Herefordshire Lagerstätte, and the evolution of the
657 holothurian body plan. Proc R Soc Lond B. 2019;286(1900):20182792.
658 46. Rouse GW, Jermiin LS, Wilson NG, Eeckhaut I, Lanterbecq D, Oji T, et al. Fixed, free, and fixed:
659 the fickle phylogeny of extant Crinoidea (Echinodermata) and their Permian–Triassic origin. Mol
660 Phylogenet Evol. 2013;66(1):161-181.

661 47. Brosseau O, Murienne J, Pichon D, Vidal N, Eléaume M, Ameziane N. Phylogeny of Cidaroida
662 (Echinodermata: Echinoidea) based on mitochondrial and nuclear markers. Org Divers Evol.
663 2012;12(2):155-165.
664 48. Durham JW, Melville R. A classification of echinoids. J Paleontol. 1957:242-272.
665 49. Suter SJ. Cladistic analysis of cassiduloid echinoids: trying to see the phylogeny for the trees. Biol
666 J Linn Soc. 1994;53(1):31-72.
667 50. Kier PM. Cassiduloids. In: Moore R, editor. Treatise on Invertebrate Paleontology, Part U.
668 Lawrence: University of Kansas Press; 1966. p. 492-523.
669 51. Mortensen T. A monograph of the Echinoidea. IV, 1 Holectypoida, Cassiduloida. Copenhagen: CA
670 Reitzel; 1948.
671 52. King B. Bayesian tip-dated phylogenetics in paleontology: Topological effects and stratigraphic
672 fit. Syst Biol. 2020;70(2):283-294.
673 53. Turner AH, Pritchard AC, Matzke NJ. Empirical and Bayesian approaches to fossil-only divergence
674 times: a study across three reptile clades. PLoS One. 2017;12(2):e0169885.
675 54. Mongiardino Koch N, Garwood RJ, Parry LA. Fossils improve phylogenetic analyses of
676 morphological characters. Proc R Soc Lond B. 2021;288(1950):2021004.
677 55. Dos Reis M, Gunnell GF, Barba-Montoya J, Wilkins A, Yang Z, Yoder AD. Using phylogenomic data
678 to explore the effects of relaxed clocks and calibration strategies on divergence time estimation:
679 primates as a test case. Syst Biol. 2018;67(4):594-615.
680 56. Warnock RC, Yang Z, Donoghue PC. Exploring uncertainty in the calibration of the molecular
681 clock. Biol Lett. 2011;8:156-159.
682 57. Dos Reis M, Thawornwattana Y, Angelis K, Telford MJ, Donoghue PC, Yang Z. Uncertainty in the
683 timing of origin of animals and the limits of precision in molecular timescales. Curr Biol.
684 2015;25(22):2939-2950.

685 58. Sauquet H. A practical guide to molecular dating. C R Palevol. 2013;12(6):355-367.
686 59. Carruthers T, Scotland RW. Uncertainty in divergence time estimation. Syst Biol. 2021;70:855-
687 861.
688 60. Carruthers T, Sanderson MJ, Scotland RW. The implications of lineage-specific rates for
689 divergence time estimation. Syst Biol. 2020;69(4):660-670.
690 61. Tao Q, Barba-Montoya J, Huuki LA, Durnan MK, Kumar S. Relative efficiencies of simple and
691 complex substitution models in estimating divergence times in phylogenomics. Mol Biol Evol.
692 2020;37(6):1819-1831.
693 62. Tao Q, Tamura K, U. Battistuzzi F, Kumar S. A machine learning method for detecting
694 autocorrelation of evolutionary rates in large phylogenies. Mol Biol Evol. 2019;36(4):811-824.
695 63. Raup DM, Sepkoski JJ. Mass extinctions in the marine fossil record. Science.
696 1982;215(4539):1501-1503.
697 64. Kier PM. Echinoids from the Triassic (St. Cassian) of Italy, their lantern supports, and a revised
698 phylogeny of Triassic echinoids. Smithson Contrib Paleobiol. 1984;56:1-41.
699 65. Smith AB. Triassic echinoids from Peru. Palaeontogr Abt A. 1994;233:177-202.
700 66. Kier PM. The poor fossil record of the regular echinoid. Paleobiology. 1977;3:168-74.
701 67. Field D, Berv J, Hsiang A, Lanfear R, Landis M, Dornburg A. Timing the extant avian radiation: The
702 rise of modern birds, and the importance of modeling molecular rate variation. Bull Am Mus Nat Hist.
703 2020;440:150-181.
704 68. Caramés A, Adamonis S, Concheyro A, Remírez M. New finding of regular echinoid elements and
705 microfossils from the Pilmatué Member, Agrio Formation (Early Cretaceous), Neuquén Basin, Argentina.
706 In: Cusminsky G, Bernasconi E, Concheyro G, editors. Advances in South American micropaleontology.
707 Springer; 2019. pp. 1-20.

708 69. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R
709 Foundation for Statistical Computing; 2019.
710 70. Ronquist F, Lartillot N, Phillips MJ. Closing the gap between rocks and clocks using total-evidence
711 dating. Phil Trans R Soc B. 2016;371(1699):20150136.
712 71. Coiro M, Doyle JA, Hilton J. How deep is the conflict between molecular and fossil evidence on
713 the age of angiosperms? New Phytol. 2019;223(1):83-99.
714 72. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
715 Bioinformatics. 2014;30(15):2114-2120.
716 73. Dunn CW, Howison M, Zapata F. Agalma: an automated phylogenomics workflow. BMC
717 Bioinformatics. 2013;14(1):330.
718 74. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length
719 transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnology.
720 2011;29(7):644-652.
721 75. Modolo L, Lerat E. UrQt: an efficient software for the Unsupervised Quality trimming of NGS
722 data. BMC Bioinformatics. 2015;16(1):1-8.
723 76. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large
724 and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31(10):1674-
725 1676.
726 77. Smit A, Hubley R, Green P. RepeatMasker Open-4.0. 2013–2015.
727 78. Hoff KJ, Stanke M. Predicting genes in single genomes with AUGUSTUS. Curr Protoc
728 Bioinformatics. 2019;65(1):e57.
729 79. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: ab initio prediction
730 of alternative transcripts. Nucleic Acids Res. 2006;34:435-439.

731 80. Simion P, Belkhir K, François C, Veyssier J, Rink JC, Manuel M, et al. A software tool
732 'CroCo'detects pervasive cross-species contamination in next generation sequencing data. BMC Biol.
733 2018;16(1):28.
734 81. Ryan JF. Alien Index: identify potential non-animal transcripts or horizontally transferred genes
735 in animal transcriptomes. 2014. doi: 10.5281/zenodo.21029.
736 82. Guang A, Howison M, Zapata F, Lawrence C, Dunn CW. Revising transcriptome assemblies with
737 phylogenetic information. PLos One. 2021;16(1):e0244202.
738 83. Ranwez V, Harispe S, Delsuc F, Douzery EJ. MACSE: Multiple Alignment of Coding SEquences
739 accounting for frameshifts and stop codons. PLoS One. 2011;6(9):e22594.
740 84. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and
741 ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007;56(4):564-577.
742 85. Morel B, Kozlov AM, Stamatakis A. ParGenes: a tool for massively parallel model selection and
743 phylogenetic tree inference on thousands of genes. Bioinformatics. 2018;35(10):1771-1773.
744 86. Mai U, Mirarab S. TreeShrink: fast and accurate detection of outlier long branches in collections
745 of phylogenetic trees. BMC Genomics. 2018;19(5):272.
746 87. Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: polynomial time species tree reconstruction
747 from partially resolved gene trees. BMC Bioinformatics. 2018;19(6):153.
748 88. Sayyari E, Mirarab S. Fast coalescent-based computation of local branch support from quartet
749 frequencies. Mol Biol Evol. 2016;33(7):1654-1668.
750 89. Aberer AJ, Kobert K, Stamatakis A. ExaBayes: massively parallel Bayesian tree inference for the
751 whole-genome era. Mol Biol Evol. 2014;31(10):2553-2556.
752 90. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic
753 algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2014;32(1):268-274.

754 91. Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A, Jermiin LS. ModelFinder: fast model
755 selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587-589.
756 92. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast
757 bootstrap approximation. Mol Biol Evol. 2017;35(2):518-522.
758 93. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-
759 friendly tool for maximum likelihood phylogenetic inference. Bioinformatics. 2019;35:4453-4455.
760 94. Wang H-C, Minh BQ, Susko E, Roger AJ. Modeling site heterogeneity with posterior mean site
761 frequency profiles accelerates accurate phylogenomic estimation. Syst Biol. 2018;67(2):216-235.
762 95. Lartillot N, Rodrigue N, Stubbs D, Richer J. PhyloBayes MPI. Phylogenetic reconstruction with
763 infinite mixtures of profiles in a parallel environment. Syst Biol. 2013;62(4):611-615.
764 96. Strimmer K, Von Haeseler A. Likelihood-mapping: a simple method to visualize phylogenetic
765 content of a sequence alignment. Proc Natl Acad Sci. 1997;94(13):6815-6819.
766 97. Shen X-X, Hittinger CT, Rokas A. Contentious relationships in phylogenomic studies can be driven
767 by a handful of genes. Nat Ecol Evol. 2017;1:0126.
768 98. Yang Z, Rannala B. Bayesian estimation of species divergence times under a molecular clock
769 using multiple fossil calibrations with soft bounds. Mol Biol Evol. 2006;23(1):212-226.
770 99. Telford MJ, Lowe CJ, Cameron CB, Ortega-Martinez O, Aronowicz J, Oliveri P, et al. Phylogenomic
771 analysis of echinoderm class relationships supports Asterozoa. Proc R Soc Lond B. 2014;281:20140479.
772 100. Schlager S. Morpho and Rvcg–Shape Analysis in R. In: Zheng G, Li S, Székely GJ, editors.
773 Statistical shape and deformation analysis. Elsevier; 2017. pp. 217-256.
774 101. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary
775 analyses in R. Bioinformatics. 2018;35(3):526-528.

Supplementary Information for:

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

Sequencing and further bioinformatic details

Apatopygus recens, *Aspidodiadema hawaiiense*, *Fellaster zelandiae*, *Histocidaris variabilis*, *Rhyncholampas pacificus*, *Sinaechinocyamus mai*, *Stereocidaris neumayeri*, *Tromikosoma hispidum*. Tissue subsamples were finely chopped with a scalpel and preserved in RNA*later* (Ambion) buffer solution for 1 day at 4°C to allow the RNAl*ater* to effectively penetrate the tissues, followed by long-term storage at -80°C until RNA extraction. Total RNA was extracted from 1.5 mm Triple-Pure High Impact Zirconium beads (Benchmark Scientific) in Trizol (Ambion), using Direct-zol RNA Miniprep Kit (Zymo Research) with in-column DNase I incubation to remove genomic DNA. Prior to mRNA capture, total RNA concentration was estimated using Qubit RNA HS Assay Kit (Invitrogen; range = 8.33-96 ng/µL), and quality was assessed using either High Sensitivity RNA ScreenTape or RNA ScreenTape with an Agilent 4200 TapeStation. Mature mRNA was isolated from total RNA and libraries were prepared using KAPA mRNA HyperPrep Kit (KAPA Biosystems) following the manufacturer's instructions (including sample customization based on total RNA quantity and quality values), targeting an insert size circa 500 base pairs (bp), and using custom 10-nucleotide Illumina TruSeq style adapters [102]. Post-amplification, DNA concentration was estimated using Qubit dsDNA BR Assay Kit (Invitrogen; range = 6.06-13.4 ng/µL). Concentration, quality, and molecular weight distribution of libraries (range = 547-978 bp) was also assessed using Genomic DNA ScreenTape with an Agilent 4200 TapeStation. Ten libraries (including two

798 annelids not employed here) were sequenced on one lane of a multiplexed run using NovaSeq S4
799 platform with 100 bp paired-end reads at the IGM Genomics Center (University of California San Diego).

800 The assembled transcriptomes were sanitized from cross-contaminant reads product of
801 multiplexed sequencing using CroCo v. 1.1 [80]. These eight datasets, as well as two annelid
802 transcriptomes not employed in this study but sequenced in the same lane, were incorporated.
803 Transcripts considered over or under-expressed across samples (as defined by default parameters),
804 were kept. This resulted in an average removal of 1.26% of assembled transcripts (range: 0.4% for
805 *Aspidodiadema* to 3.06% for *Histocidaris*).

806

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

816

817 *Bathysalenia phoinissa*, *Micropyga tuberculata*: The material was collected by the deep-sea
818 cruise BIOMAGLO [103] conducted in 2017 jointly by the French National Museum of Natural History
819 (MNHN) as part of the Tropical Deep-Sea Benthos program, the French Research Institute for
820 Exploitation of the Sea (IFREMER), the "Terres Australes et Antarctiques Françaises" (TAAF), the

821 Departmental Council of Mayotte and the French Development Agency (AFD), with the financial support
822 of the European Union (Xe FED). Specimens were sorted into taxonomic bins at class level and
823 collectively fixed in 70% ethanol at room temperature. Following taxonomic identification, tissue
824 samples were obtained using sterilized forceps and disposable scalpel blades. Tissue subsamples were
825 collected in high-purity 96% ethanol and stored at -40°C until DNA extraction. Total genomic DNA was
826 extracted using the DNeasy Blood and Tissue Kit (QIAGEN) following the manufacturer's instructions,
827 complemented by the addition of 4 µl RNAse A (QIAGEN, 100 mg/ml) for RNA digestion. Elution buffer
828 volume was reduced to 60 µl and the elution step repeated twice per sample in order to maximize DNA
829 concentration and yield. After collection of the first elution, spin-column membranes were rinsed again
830 twice with fresh elution buffer (40 µl) to further increase yield. Elutions were collected separately, with
831 the first one used for sequencing and the second for quality control. For the latter, a portion of the
832 mitochondrial cytochrome c oxidase subunit I (COI) was amplified using PCR. Amplifications were
833 conducted with TopTaq DNA Polymerase (QIAGEN) using 1 µl of extracted genomic DNA (approx. 10–15
834 ng). Details on primers and protocols can be found in [104]. PCR products were visualized on a 1%
835 agarose gel, purified using ExoSAP-IT (Affymetrix), and sequenced at Microsynth GmbH (Vienna, Austria)
836 with the same primers. Sequences are deposited in GenBank under the accession numbers XXX and XXX.

837 Prior to library preparation, total genomic DNA concentration was estimated using Qubit DNA
838 BR Assay Kit (Invitrogen) in order to determine library preparation strategy. This step was carried out by
839 Macrogen (Korea), using an Illumina TruSeq DNA PCR-free kit (for *Micropyga*) and an Illumina TruSeq
840 Nano DNA kit (for *Bathysalenia*). Library preparation followed the manufacturer's instructions, targeting
841 an insert size of 350 base pairs (bp). The libraries were sequenced by Macrogen (Korea) on a shared lane
842 of a multiplexed run using an Illumina HiSeq X instrument with 150 bp paired-end reads. Sequencing
843 depth was 134.1 and 161.3 million reads for *Bathysalenia* and *Micropyga*, respectively. Pre-processing of
844 Illumina shotgun data was carried out by removal of adapters using BBDuk

845 (https://sourceforge.net/projects/bbmap/) with settings: `minlen=50 ktrim=r k=23 mink=11`
846 `tpe tbo`); followed by quality trimming of reads using UrQt v. 1.0.18 [75], discarding regions with
847 phred quality score < 28 and sequences of size < 50. Assembly of the data was carried out with MEGAHIT
848 v. 1.1.2 [76], in an iterative approach followed by assessment of assembly quality using BUSCO v. 3.0.2.
849 scores [105] using the metazoan dataset ODB v. 9 [106]. Final BUSCO scores of the draft assemblies
850 were 28.4% (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%,
851 F:59.0%, M:23.8%, n:978) for *Micropyga*. As recommended by Hoff and Stanke [78], draft genomes were
852 masked with RepeatMasker v. 4.1.0 [77] prior to gene prediction, using settings: `-norna -xsmall`;
853 resulting in 3.82 % and 2.9% masked bases for *Bathysalenia* and *Micropyga*, respectively. Gene
854 prediction was carried out using AUGUSTUS v. 3.2.3 [79] using settings: `--strand=both --`
855 `singlestrand=false --genemodel=partial --codingseq=on --sample=0 --`
856 `alternatives-from-sampling=false --exonnames=on --softmasking=on`. This step
857 relied on a training dataset composed of universal single-copy orthologs (USCOs) derived from the most
858 recent *Strongylocentrotus purpuratus* genome (Spur v.5).

859

860 *Echinothrix calamaris*, *Tripneustes gratilla*. Tissue samples were taken from live sea urchins
861 housed in laboratory aquaria and total RNA was immediately extracted from 150 mg of tissue using a
862 PureLink™ RNA extraction kit (Invitrogen™). The quality of total RNA was assessed on a BioAnalyzer
863 2100 (Agilent Technologies, Santa Clara, CA, USA) to ensure a RIN > 9 for all samples. Mature mRNA was
864 extracted from 1ug of total RNA and cDNA libraries were constructed using a TruSeq kit (Illumina).
865 Quality control of libraries was assessed on a BioAnalyzer and quantification measured using qPCR.
866 *NEBNext* Multiplex adaptors were added via ligation, and the cDNA libraries were sequenced at Genome
867 Quebec, McGill University. Three libraries were multiplex sequenced on one lane of Illumina at a

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

870

871 *Tetrapygus niger*. Adult specimens were collected and transported alive to the University of
872 Concepcion. Tissue samples (female gonad and tube feet) from one individual were finely chopped with
873 a scalpel, and total RNA was extracted immediately using TriReagent (Sigma), following the
874 manufacturer's instructions. Total RNA concentration was estimated using QuantiFluor® RNA System
875 (111.84-339.04 ng/µL) and quality was assessed using Agilent RNA 6000 Nano kit with an Agilent 2100
876 TapeStation (RIN: 5.4-8.4). cDNA libraries were prepared at Genoma Mayor SpA (Chile) using TruSeq
877 Stranded mRNA Kit (Illumina) and sequenced on a multiplexed run using Illumina HiSeq 4000 platform
878 with 150 paired-end reads, resulting in 54.1 M reads for gonad female and 52.7 M reads for tube feet,
879 respectively. These two transcriptomic datasets were combined before performing all steps of
880 bioinformatic processing.

881

882 *Eucidaris metularia*. A single specimen was preserved in RNA*later* after being starved overnight
883 and kept for 1 day at 4°C before long-term storage at -80°C. Total RNA was extracted using Direct-zol
884 RNA Miniprep Kit (with in-column DNase treatment; Zymo Research) from Trizol. mRNA was isolated
885 with Dynabeads mRNA Direct Micro Kit (Invitrogen). RNA concentration was estimated using Qubit RNA
886 broad range assay kit, and quality was assessed using RNA ScreenTape with an Agilent 4200 TapeStation.
887 Values were used to customize downstream protocols following manufacturers' instructions. Library
888 preparation was performed with KAPA-Stranded RNA-Seq kits, targeting an insert size in the range of
889 200–300 base pairs. The quality and concentration of libraries were assessed using DNA ScreenTape.
890 The library was sequenced in a multiplexed pair-end runs using Illumina HiSeq 4000 with 7 other

891 libraries in the same lane (all previously published in [4]). In order to minimize read crossover, 10 bp
892 sequence tags designed to be robust to indel and substitution errors were employed [102].

893 The assembled transcriptome was filtered from cross-contaminant reads product of multiplexed
894 sequencing using CroCo v. 1.1 [80]. Seven other transcriptomes sequenced together were also
895 employed, and results of this step are reported in figure S1 of Mongiardino Koch et al. [3].

896

Fossil constraints

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

Ambulacraria (Echinodermata-Hemichordata divergence) – **Younger bound**: 518 Ma, Middle Atdabanian (age 3, Series 3 of Cambrian). **Older bound**: 636.1 Ma, Lantian Biota, Ediacaran. **Reference:** [107]. **Notes:** The root of our tree is constrained with a younger bound concurrent with the earliest occurrence of stereom in the fossil record (see below). The older bound is based on the maximum age of the Lantian Biota, a Lagerstätte lacking any trace of eumetazoan fossils.

Echinodermata (Pelmatozoan-Eleutherozoan divergence) – **Younger bound**: Unconstrained. **Older bound:** 515.5 Ma, Middle Atdabanian (age 3, Series 3 of Cambrian). **Reference**: [107]. **Notes:** This divergence (which represents the divergence of crown group echinoderms), has an older bound based upon the earliest occurrence of stereom in the fossil record. Stereom constitutes an echinoderm synapomorphy readily recognizable in the fossil record due to its unique mesh-like structure [41]. The first records of stereom point to a sudden and global appearance within the middle Atdabanian [40]. Note that this is also used as younger bound for the divergence between echinoderms and hemichordates.

Echinozoa (Echinoidea-Holothuroidea divergence) – **Younger bound:** 469.96 Ma, Top of *Pseudoclamacograptus decorates* graptolite zone. **Older bound:** 461.95 Ma, Top Floian. **Reference:** [108, 109]. **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**: [108, 109]. **Taxon**:

919 *Maydenia roadsidensis.* **Notes**: The divergence of asterozoan classes is calibrated based upon the
920 stratigraphic occurrence of *Maydenia roadsidensis*, the oldest asterozoan, which is a stem group
921 ophiuroid [110] and occurs in the *Psigraptus jacksoni* Zone of the Floian.

922 Asteroidea – **Younger bound:** 239.10 Ma, Top Fassanian, Ladinian, Mo3, *nodosus* biozone, Triassic.
923 **Older bound**: 252.16 Ma, P-T Boundary. **Reference**: [111]. **Taxon**: *Trichasteropsis weissmanni*. **Notes**:
924 Fossil evidence suggests the divergence of the asterozoan crown group took place sometime in the Early
925 or Middle Triassic [111]. Based upon its phylogenetic position and stratigraphic occurrence as the oldest
926 crown group asteroid, we use the forcipulatacean *T. weissmanni* from the Middle Triassic Muschelkalk
927 as the soft bound on this divergence.

928 Crinoidea – **Younger bound**: 247.06 Ma, Top Spathian (Paris Biota, Lower Shale unit, Thaynes Group,
929 Spathian, Triassic). **Older bound**: Unconstrained. **Reference**: [112]. **Taxon**: *Holocrinus*. **Notes**: Crown
930 group Crinoidea is difficult to define, though a rapid post-Palaeozoic morphological diversification is
931 supported by fossil evidence. We use *Holocrinus* from the Early Triassic of the Thaynes group to calibrate
932 the younger bound on the divergence of the crown group (i.e., the split between Isocrinida and all other
933 extant crinoids).

934 Ophiuroidea (Euryophiurida-Ophintegrida divergence) – **Younger bound**: 247.06 Ma, Top Spathian (Paris
935 Biota, Lower Shale unit, Thaynes Group, Spathian, Triassic). **Older bound**: Unconstrained. **Reference**:
936 [113]. **Taxon**: *Shoshonura brayardi*. **Notes**: *S. brayardi* from the Spathian Thaynes Group of the Western
937 USA is a member of the crown group ophiuroid clade Ophiodermatina. It is thus the currently oldest
938 described member of the ophiuroid crown group, setting its minimum age of divergence.

939 Pneumonophora (Holothuriida-Neoholothuriida divergence) – **Younger bound**: 259.8 Ma, S*pinosus* zone
940 of Early Ladinian (used base Ladinian). **Older bound**: 241.5 Ma, Base Wuchiapingian. **Reference**: [44,

108]. **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**: [114]. **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. [115]. **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.

963 Stomopneustoida-Camarodonta divergence – **Younger bound**: 201.3 Ma, Top Pliensbachian. **Older**
964 **bound**: 182.7 Ma, Base Jurassic. **Reference**: See discussion in supplement of Thompson et al. [13].
965 **Taxon**: Stomechinids from Morocco like *Diplechinus hebbriensis* [116]. **Notes**: The oldest stomechinids,
966 which are stem group stomopneustoids, are from the Early Jurassic of Morocco.

967 Neognathostomata-Atelostomata divergence – **Younger bound**: 234.5 Ma, Base of Toarcian. **Older**
968 **bound**: 182.7 Ma, St. Cassian Beds, bottom of Julian 2 ammonoid Zone. **Reference**: [14, 115]. **Taxon**:
969 Younger bound is set by *Galeropygus sublaevis* (older bound is relatively uninformative). **Notes**: The
970 base of the Toarcian (which contains the oldest neognathostomate *G. sublaevis*) is used as the lower
971 bound on the divergence between the neognathostomata and all other crown irregular echinoids. Note
972 that depending on the position of the unsampled echinoneoids, this split might also correspond to the
973 origin of crown group Irregularia (see recent topologies recovered by Lin et al. [36] and Mongiardino
974 Koch & Thompson [3]).

975 Atelostomata (Holasteroida-Spatangoida divergence) – **Younger bound**: 137.68 Ma, Bottom of
976 *Campylotoxus* zone in Valanginian. **Older bound**: 201.3 Ma, Base of Jurassic. **Reference**: [117]. **Taxon**:
977 Younger bound is based on *Toxaster* (older bound is loosely uninformative). **Notes**: The toxasterids are
978 stem group spatangoids that calibrate the divergence between spatangoids and holasteroids. The oldest
979 *Toxaster* are from the *Campylotoxus* zone in the Valanginian, and set the younger bound on the
980 divergence.

981 Echinolampadacea (Cassiduloida-Scutelloida divergence) – **Younger bound**: 113 Ma, Base Albian. **Older**
982 **bound**: 145 Ma, Base Cretaceous. **Reference**: [118]. **Taxon**: Younger bound is based on *Eurypetalum*
983 rancheriana (older bound is loosely uninformative). **Notes**: The oldest member of Echinolampadacea is
984 the faujasiid *Eurypetalum rancheriana* (originally placed in the genus *Faujasia* and later transferred, see
985 [28, 29]) from the Late Albian of Colombia.

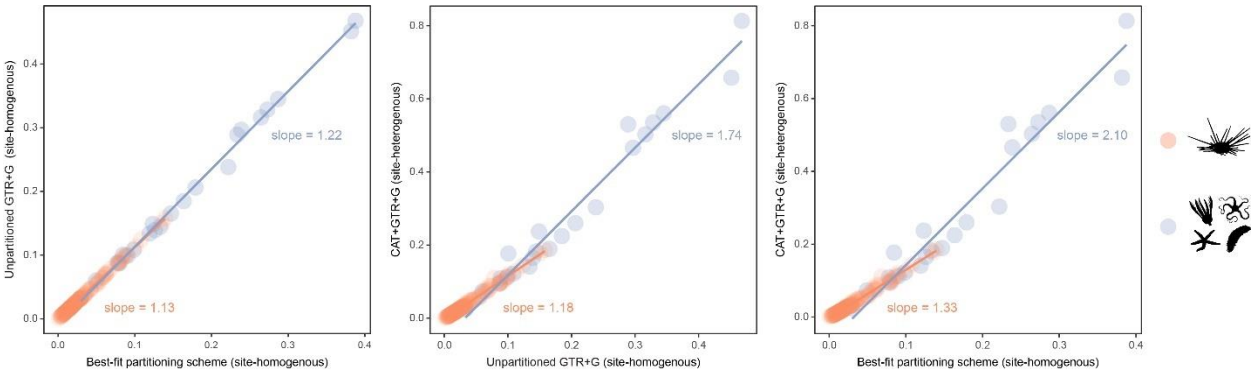
986 Scutelloida (Scutelliformes-Laganiformes divergence) – **Younger bound**: 56.0 Ma, Bottom of Eocene.
987 **Older bound**: Unconstrained. **Reference**: [14, 119]. **Taxon**: Younger bound is based on *Eoscutum*
988 *doncieuxi*. **Notes**: We use the base of the Eocene as the younger bound on the divergence between
989 scutelliforms and laganiforms. This is based upon the occurrence of *E. doncieuxi*, the oldest scutelliform,
990 which is known from the Lower Eocene. The older bound is left unconstrained in order to account for
991 the current uncertainty in the origin of the clade (i.e., allow for Mesozoic ages).

992 *Leodia*-(*Encope*+*Mellita*) divergence – **Younger bound**: 23.03 Ma, Base Miocene. **Older bound**: 33.9 Ma,
993 Base of Oligocene. **Reference**: [120, 121]. **Taxon**: Younger bound is based on *Encope michoacanensis*
994 (older bound is loosely uninformative). **Notes**: This node is constrained using the oldest representative
995 of these three genera of mellitids.

996 Clypeasteroida – **Younger bound**: 47.8 Ma, Base Lutetian. **Older bound**: Unconstrained. **Reference**:
997 [122, 123]. **Taxon**: Younger bound is based on *Clypeaster calzadai* and *Clypeaster moianensis*. **Notes**:
998 The oldest known fossil species of Clypeasteroida (the clypeasterids *C. calzadai* and *C. moianensis*) are
999 from the middle Eocene, upper Lutetian (Biarritzian) of Cataluña, Spain. The older bound is left
1000 unconstrained in order to account for the current uncertainty in the origin of the clade (i.e., allow for
1001 Mesozoic ages).

1002 Strongylocentrotidae-Toxopneustidae divergence – **Younger bound**: 44.4 Ma, Eocene Planktonic Zones
1003 E9-E12. **Older bound**: 56.0 Ma, Base Eocene. **Reference**: [124, 125]. **Taxon**: Younger bound is based on
1004 *Lytechinus axiologus* (older bound is loosely uninformative). **Notes**: The oldest member of the clade
1005 comprising toxopneustids and strongylocentrotids is the toxopneustid *Lytechinus axiologus.* This taxon is
1006 known from the Eocene Yellow Limestone Group of Jamaica. The Yellow Limestone group represents the
1007 Eocene Planktonic Zones E9-E12.

1008 **Supplementary Figures and Tables**

1009 

1010 **Fig. S1**: Estimated branch lengths across different models of molecular evolution. Different site-
1011 homogeneous models (left) infer similar levels of divergence, and the choice between them induces
1012 little distortion in the general tree structure. Site-heterogeneous models on the other hand, not only
1013 infer a larger degree of divergence between terminals relative to site-homogeneous ones (center and
1014 left), but they also distort the tree (i.e., impose a non-isometric stretching), with branch lengths
1015 connecting outgroup taxa expanding much more than those within the ingroup clade.

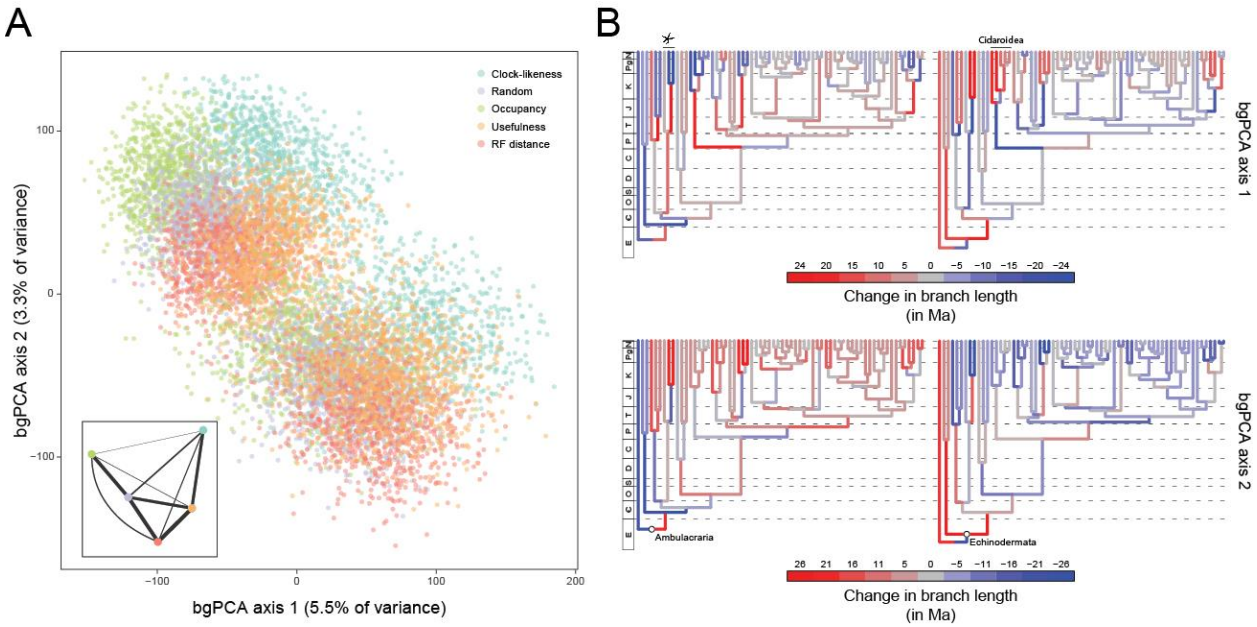


Fig. S2: Effect of gene sampling strategy on inferred times of divergence. **A.** A chronospace, depicting the differences in inferred ages between chronograms obtained using different loci. Axes constitute the first two dimensions of the bgPCA (capturing 8.8% out of a total of 10.5% of variances summarized by all four dimensions). Note that a bimodal distribution of the data, likely corresponding to the implementation of LN or UGAM clock models (see Fig. 3) is evident, even when the bgPCA is performed so as to discriminate chronograms based on a different factor. The inset shows the centroid for each type of loci sampling strategy, 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.

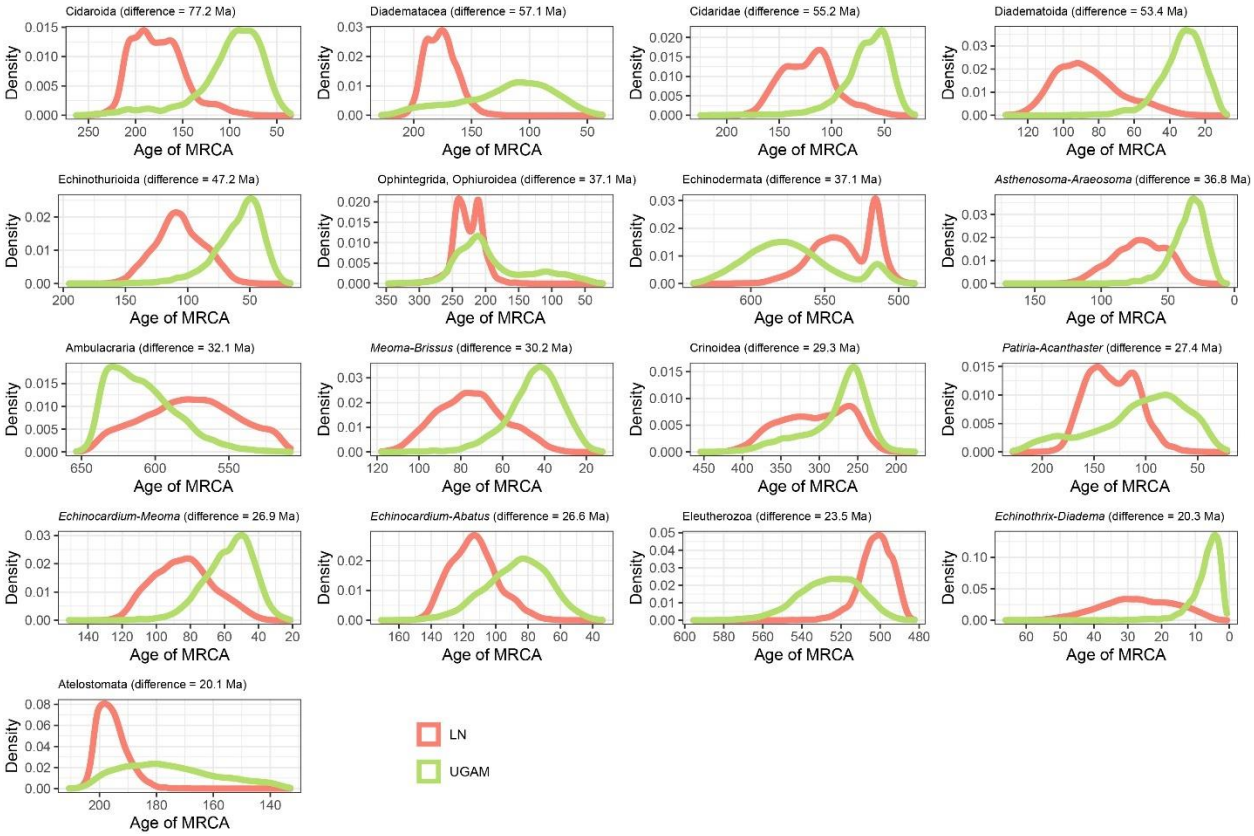


Fig. S3: Distribution of posterior probabilities for node ages that show an average difference larger than 20 Ma depending on the choice of clock prior.

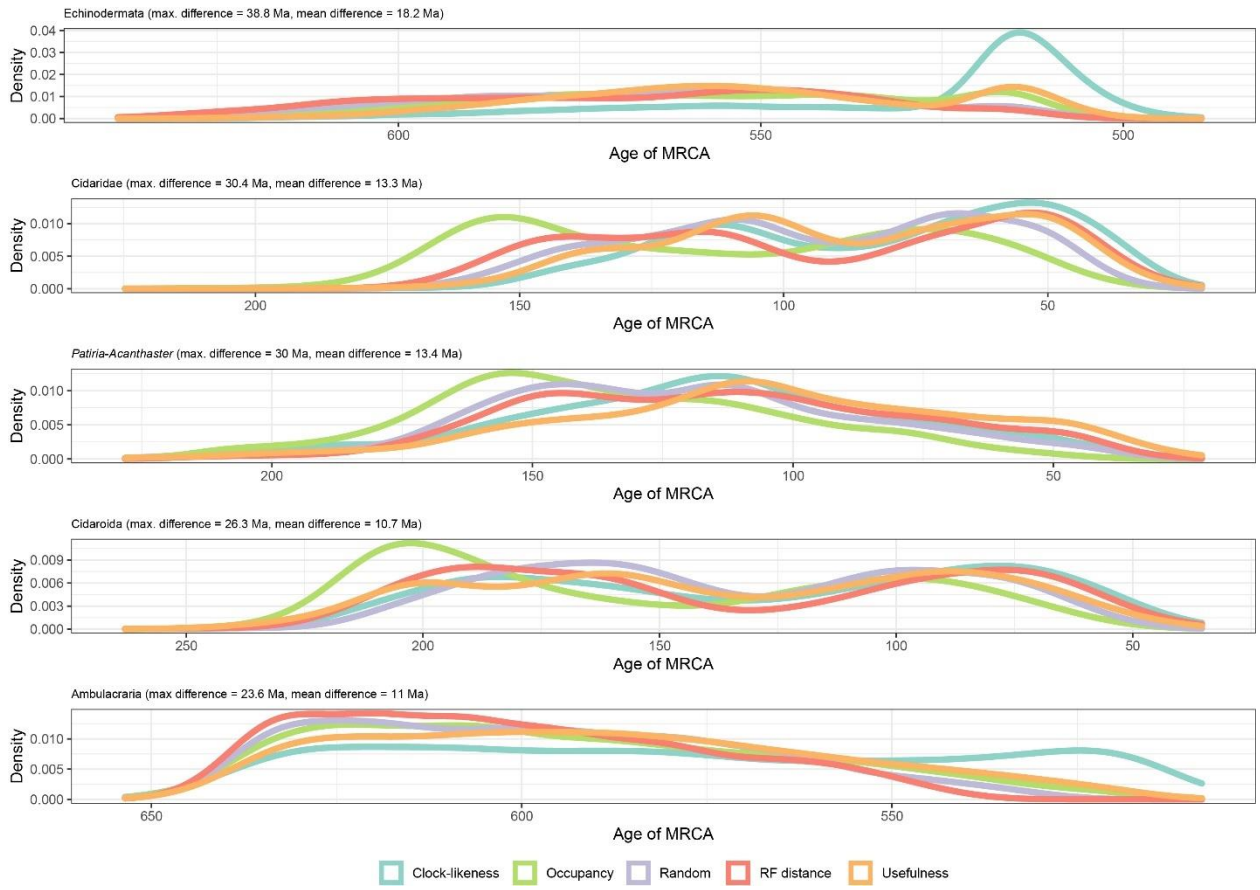


Fig. S4: Distribution of posterior probabilities for node ages that show a maximum difference larger than 20 Ma depending on the choice of clock prior (no node showed average differences larger than 20 Ma). The largest differences can be seen in the relatively younger ages of Ambulacraria and Echinodermata 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 Fig. S2 as short distances between their centroids in chronospace.

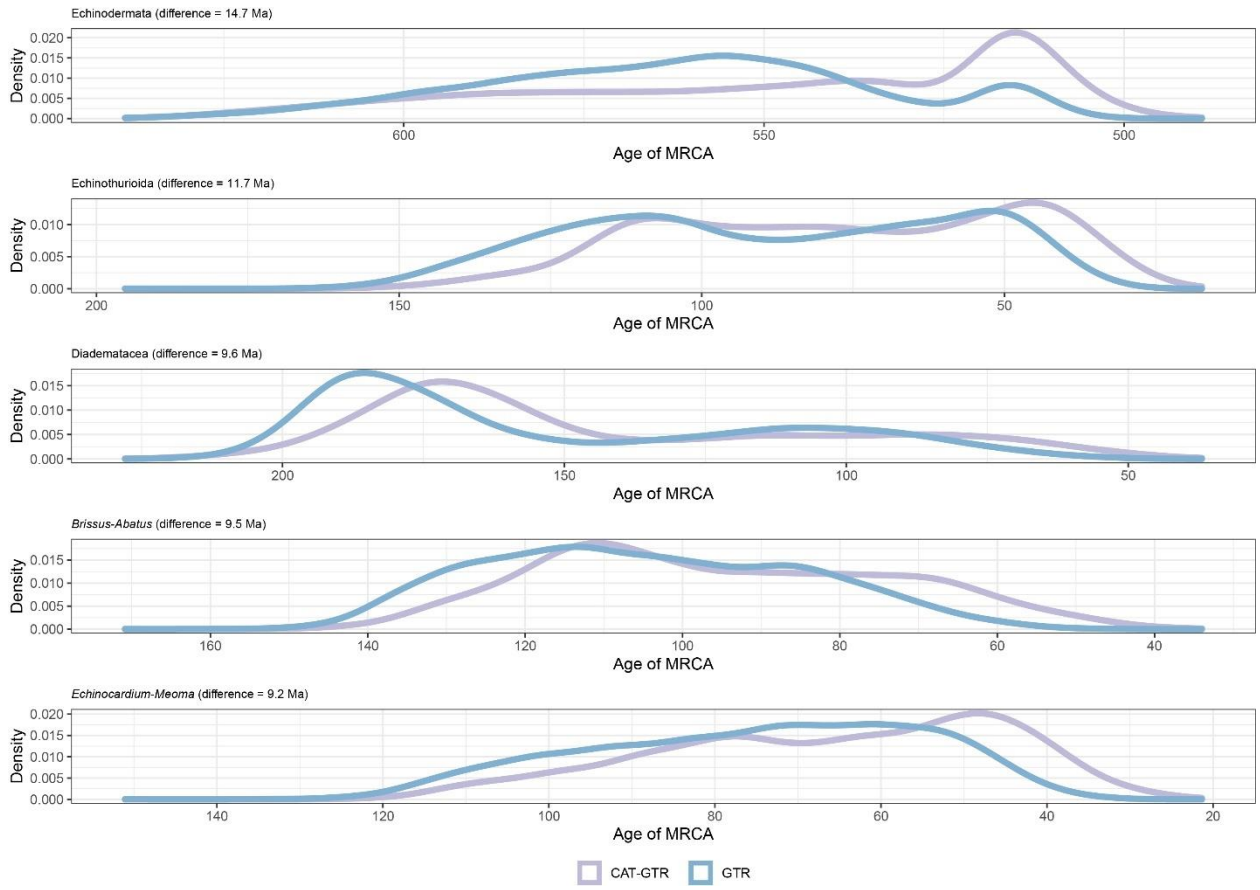
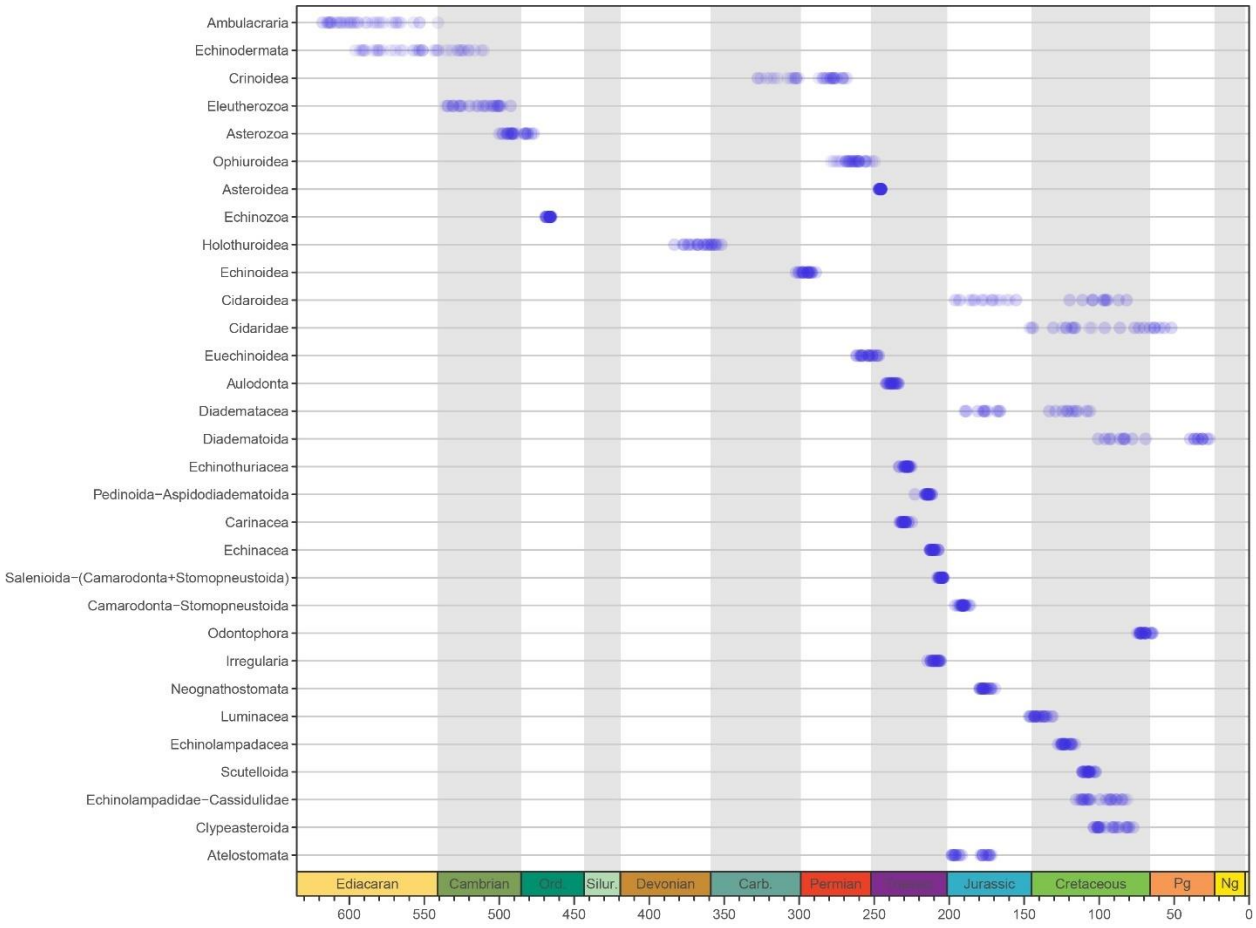


Fig. S5: Distribution of posterior probabilities for node ages that are the most affected by the choice of model of molecular evolution. No node showed average differences larger than 20 Ma, so those with the largest effect are shown.



1043

1044 **Fig. S6:** Median ages for selected clades across the consensus trees of the 40 time-calibrated

1045 experiments performed (20 analyses x 2 runs each).

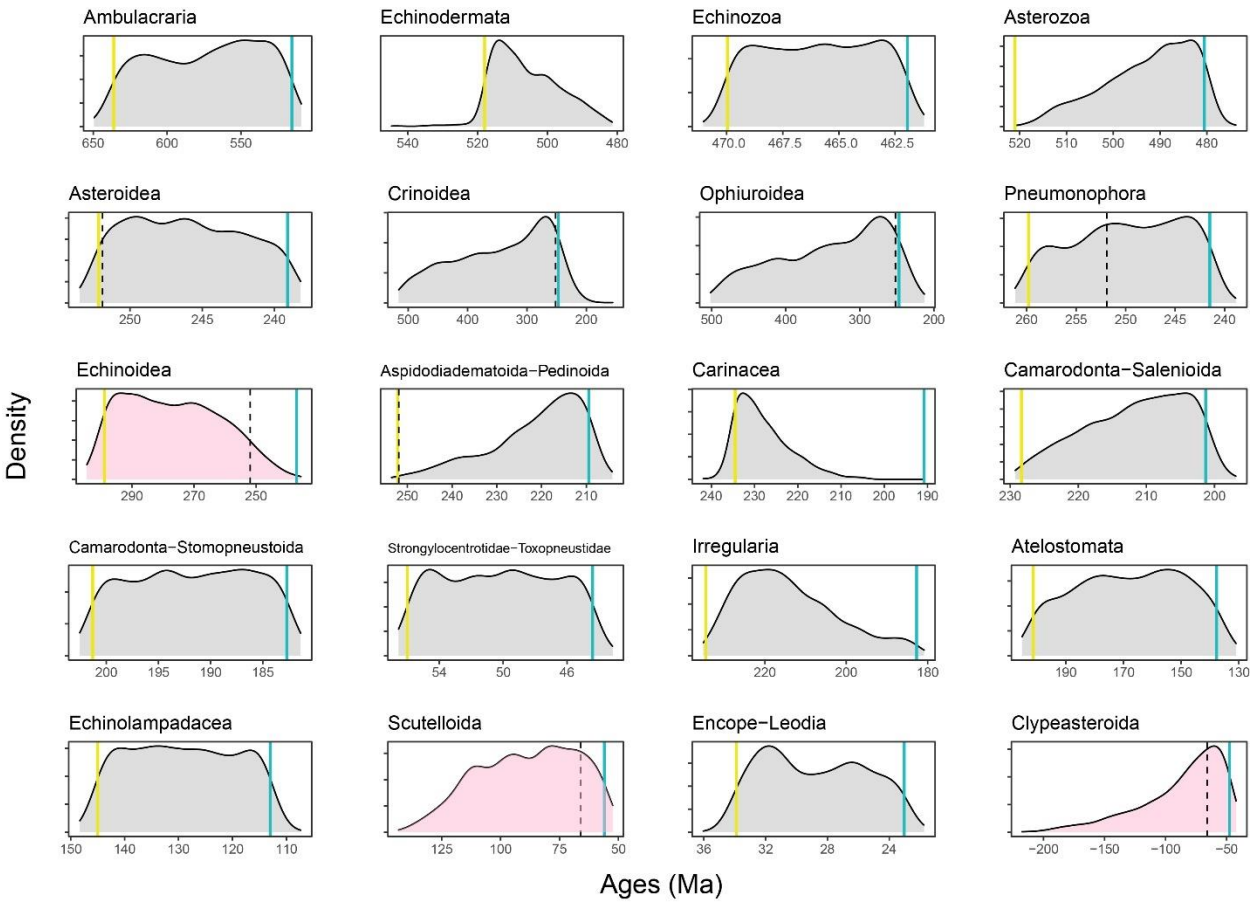


Fig. S7: Prior distributions of all constrained nodes. Five hundred replicates were sampled from the joint prior, showing appropriately broad distributions of node ages. Blue lines show minima and yellow ones maxima (when enforced); dotted lines show the age of the Permian-Triassic (251.9 Ma) and Cretaceous-Paleogene (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.

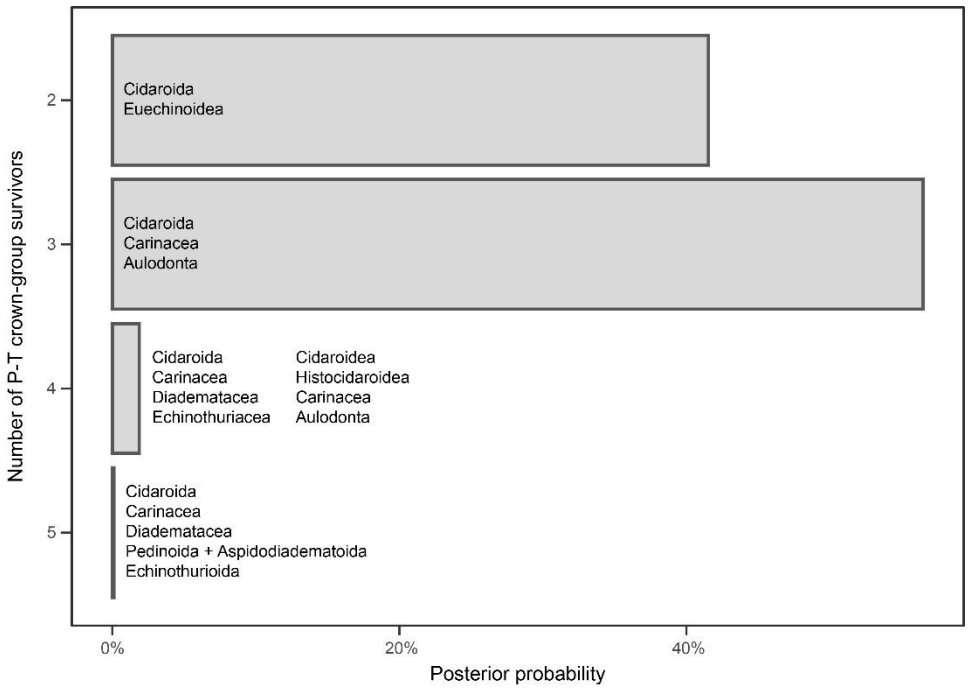


Fig. S8: Number of lineages inferred to have crossed the Permian-Triassic (P-T) boundary. The probabilities of each scenario are estimated from the inferred divergence times of the 10,000 chronograms sampled across all of the analyses performed (250 for the two runs of each combination of sampled loci, model of molecular evolution, and clock prior). The probability of three or more crown group lineages surviving the Permian-Triassic extinction is 58.47%.

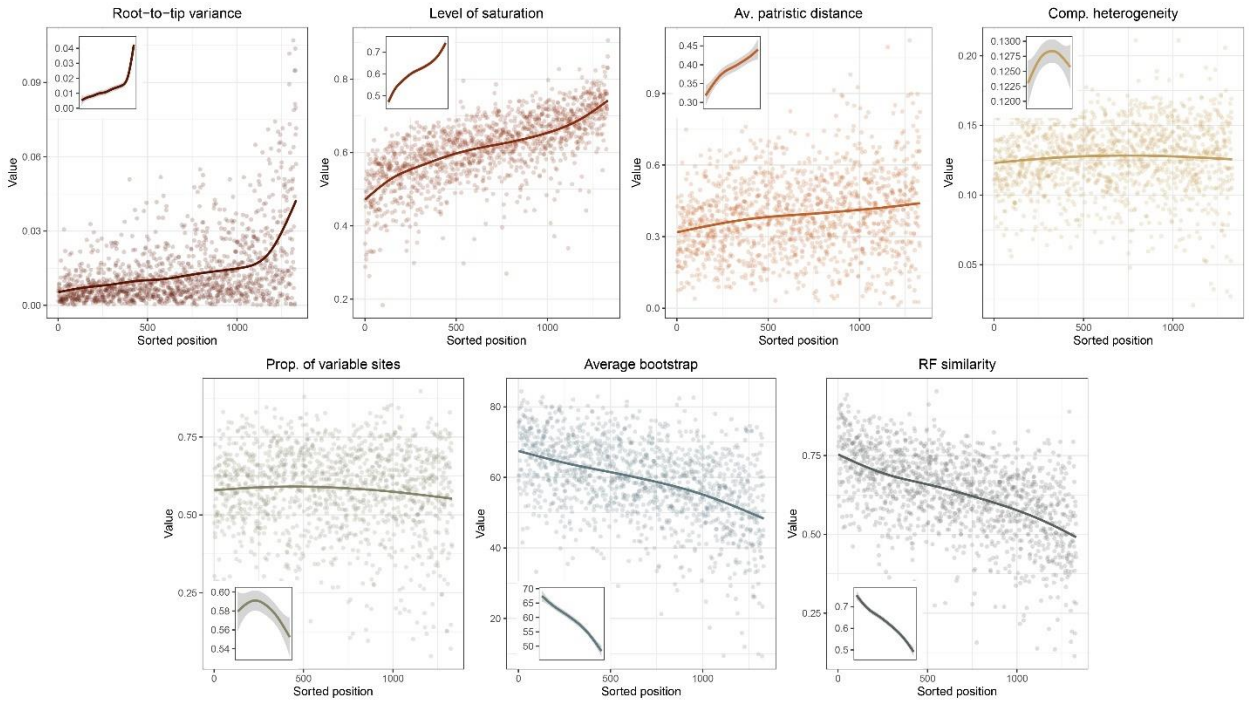


Fig. S9: Ordering of loci enforced using *genesortR* [37] and its relationship to the seven gene properties employed. High ranking loci (i.e., the most phylogenetically useful) show low root-to-tip variances (or high clock-likeness), low saturation and compositional heterogeneity, as well as high average bootstrap and Robinson-Foulds similarity to a target topology (in this case, with the contentious relationship among major lineages of Echinacea collapsed).

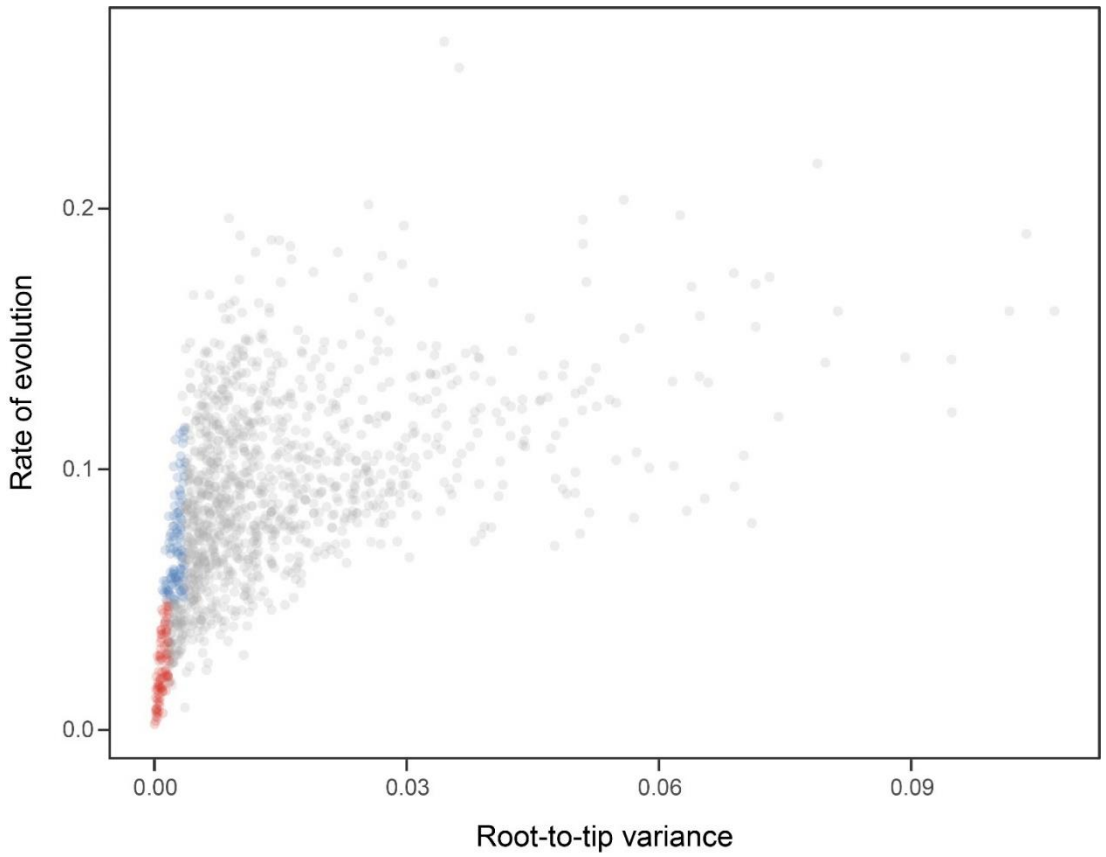


Fig. S10: Relationship between the root-to-tip variance (a proxy for the clock-likeness of loci) and the rate of evolution (estimated as the total tree length divided by the number of terminals). The most clock-like loci (shown in red), which are often favored for the inference of divergence times (e.g., [60, 126]), are among the most highly conserved and can provide little information for constraining node ages [37]. Highly clock-like genes with a higher information content were used instead by choosing the loci with the lowest root-to-tip variance from among those that were within 1 standard deviation from the mean evolutionary rate (shown in blue).

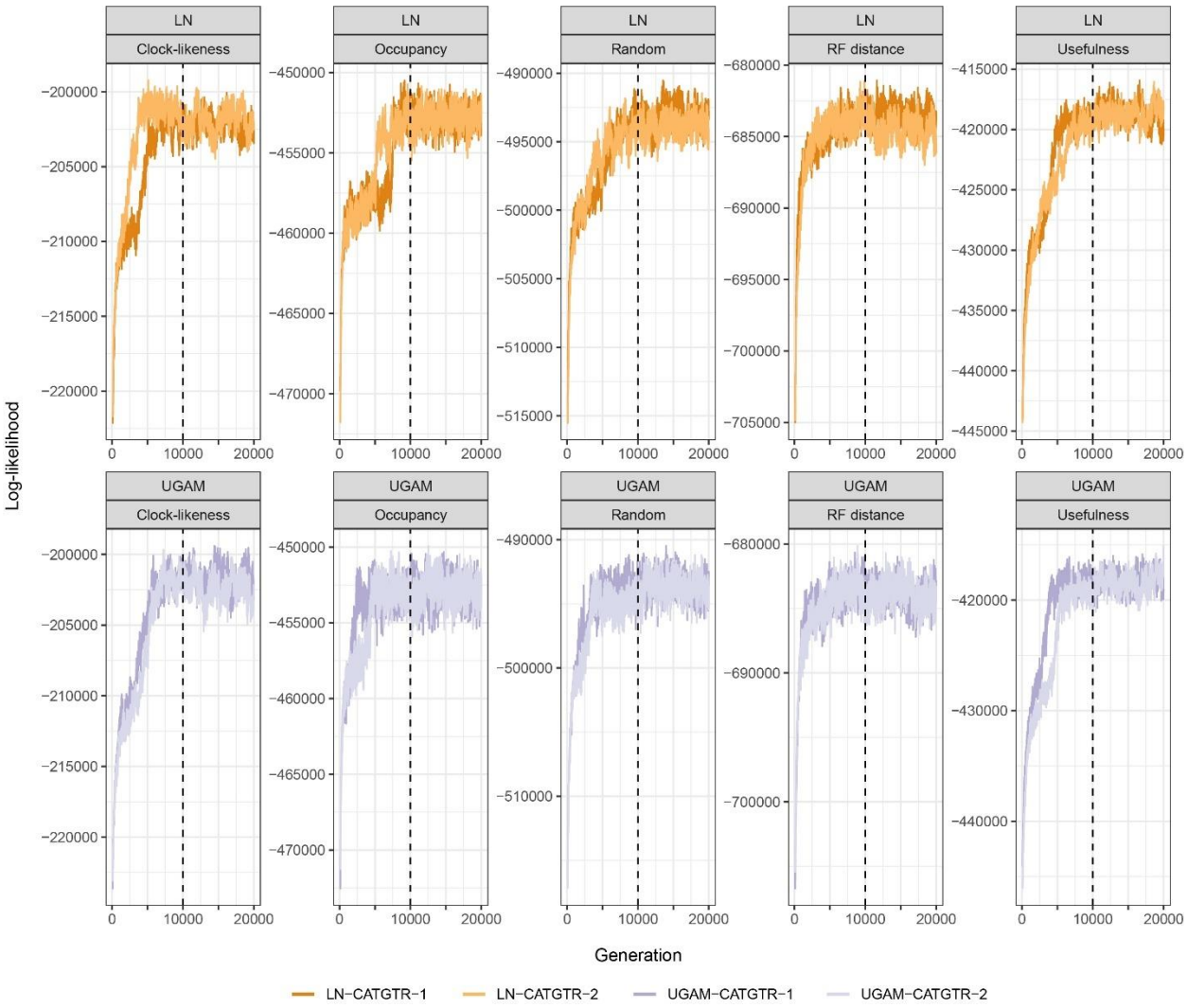


Fig. S11: Trace plots of the log-likelihood of different time-calibration runs. All runs show evidence of reaching convergence and stationarity before our imposed burn-in fraction of 10,000 generations. For simplicity, only runs under the CAT+GTR+G model are plotted. Those run under GTR+G converged much faster.

1078 **Table S1:** Transcriptomic/genomic datasets added in this study relative to the taxon sampling of
1079 Mongiardino Koch et al. [4] and Mongiardino Koch & Thompson [3]. Taxa with citations were taken from
1080 the literature, and details can be found in the corresponding studies and associated NCBI BioProjects.
1081 Taxa are shown in alphabetical order; those identified with "OG" are outgroup taxa (i.e., non-echinoids).
1082 Voucher specimens are deposited at the Benthic Invertebrate Collection, Scripps Institution of
1083 Oceanography (SIO-BIC), and the Echinoderm Collection, Muséum National d'Histoire Naturelle (MNHN-
1084 IE). If multiple accession numbers are shown for a given taxon, these datasets were merged before
1085 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)	[127]	-	-	-	-	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)	[128]	-	-	-	-	SRX445860
Astrophyton muricatum (OG)	[129, 130]	-	-	-	-	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)	[130, 131]	-	-	-	-	SRX1397823
Leodia sexiesperforata	This study	Eggs	Frances Armstrong	Bocas del Toro, Panama	-	
Metacrinus rotundus (OG)	[132]	-	-	-	-	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)	[128]	-	-	-	-	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	[133]	-	-	-	-	DRX187887 DRX187888
Sinaechinocyamus mai	This study	Gregory's diverticulum	Jih-Pai Lin	Miaoli County, Taiwan	SIO-BIC E7919	
Sterechinus neumayeri	[134]	-	-	-	-	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

1086

1087 **Table S2:** Details of molecular datasets and supermatrix. Statistics for raw reads and assemblies are
1088 shown for datasets incorporated in this study relative to the sampling of [3, 4] (where similar statistics
1089 can be found for the other datasets). Taxa are shown in alphabetical order; those identified with "OG"
1090 are outgroup taxa (i.e., non-echinoids). Novel datasets correspond to Illumina pair-end transcriptomes,
1091 except for two draft genomes (*Bathysalenia phoinissa* and *Micropyga tuberculata*). Mean insert size is
1092 expressed in number of base pairs. For transcriptomes, read pairs (initial) shows numbers input into
1093 Agalma [73], (i.e., after processing with Trimmomatic [72] or UrQt [75]), while read pairs (retained)
1094 shows those that passed the Agalma curation checks (including ribosomal, adapter, quality, and
1095 compositional filters), and represent the final number of read pairs used for assembly. For genomes, see
1096 information in the bioinformatic details above. Assemblies were further sanitized with either
1097 alien_index [81] alone or in combination with CroCo [80] (see bioinformatic details), and the number of
1098 assembled transcripts/contigs shows the size of datasets after these curation steps. The number of loci
1099 shows the occupancy of terminals in the supermatrix output by Agalma (1,346 loci at 70% occupancy),
1100 after which loci were further removed by TreeShrink [86], 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

1101

1102 **Supplementary References (references 1-101 in main text)**

1103 102. Faircloth BC, Glenn TC. Not all sequence tags are created equal: designing and validating
1104 sequence identification tags robust to indels. PLoS One. 2012;7(8):e42543.
1105 103. BIOMAGLO cruise, RV Antea. 2017. doi: 10.17600/17004000.
1106 104. Bronstein O, Kroh A, Miskelly AD, Smith SD, Dworjanyn SA, Mos B, et al. Implications of range
1107 overlap in the commercially important pan-tropical sea urchin genus *Tripneustes* (Echinoidea:
1108 Toxopneustidae). Marine Biology. 2019;166(3):34.
1109 105. Seppey M, Manni M, Zdobnov EM. BUSCO: assessing genome assembly and annotation
1110 completeness. Methods Mol Biol. 2019;1962:227-245.
1111 106. Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simao FA, Pozdnyakov IA, et al. OrthoDB
1112 v8: update of the hierarchical catalog of orthologs and the underlying free software. Nucleic Acids Res.
1113 2015;43:250-256.
1114 107. Benton MJ, Donoghue PC, Asher RJ, Friedman M, Near TJ, Vinther J. Constraints on the timescale
1115 of animal evolutionary history. Palaeontol Electron. 2015;18(1):1-106.
1116 108. Erkenbrack EM, Thompson JR. Cell type phylogenetics informs the evolutionary origin of
1117 echinoderm larval skeletogenic cell identity. Comm Biol. 2019;2(1):160.
1118 109. Cooper R, Sadler P, Hammer O, Gradstein F. The Ordovician Period. In: Gradstein FM, Ogg JG,
1119 Schmitz M, Ogg G, editors. The geologic time scale 2012. Elsevier; 2012. pp. 489-523.
1120 110. Hunter AW, Ortega-Hernández J. A new somasteroid from the Fezouata Lagerstätte in Morocco
1121 and the Early Ordovician origin of Asterozoa. Biol Lett. 2021;17(1):20200809.
1122 111. Villier L, Brayard A, Bylund KG, Jenks JF, Escarguel G, Olivier N, et al. *Superstesaster promissor*
1123 gen. et sp. nov., a new starfish (Echinodermata, Asteroidea) from the Early Triassic of Utah, USA, filling a
1124 major gap in the phylogeny of asteroids. J Syst Palaeontol. 2018;16(5):395-415.

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

113. Thuy B, Escarguel G. A new brittle star (Ophiuroidea: Ophiodermatina) from the Early Triassic Paris Biota (Bear Lake County, Idaho, USA). Geobios. 2019;54:55-61.

114. Kroh A. Echinoids from the Triassic of St. Cassian-A review. Geo Alp. 2011;8:136-140.

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

116. Smith AB. *Gymnodiadema* and the Jurassic roots of the Arbacioida (stirodont echinoids). Swiss J Palaeontol. 2011;130(1):155-171.

117. François É, Marchand D, David B. Fluctuations morphologiques et hétérochronies chez *Toxaster* (échinides, Crétacé inférieur). C R Palevol. 2003;2(6-7):597-605.

118. Cooke CW. Some Cretaceous echinoids from the Americas: US Geol Surv Prof Pap. 1955;264E:87-112.

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

120. Durham JW. Fossil *Encope* (Echinoidea) from the Pacific coast of southern Mexico. Rev Mex Cienc Geol. 1994;11(1):113-116.

121. Coppard SE, Lessios HA. Phylogeography of the sand dollar genus *Encope*: implications regarding the Central American Isthmus and rates of molecular evolution. Sci Rep. 2017;7(1):11520.

122. Ali MSM. The paleogeographic distribution of *Clypeaster* (Echinoidea) during the Cenozoic Era. Neues Jahrb Geol Paläontol. 1983;8:449-464.

1149 123. Via L, Padreny J. Dos nuevas especies de *Clypeaster* del Eoceno de Cataluña. Publ Inst Invest
1150 Geol Diput Prov. 1970;24:89-118.
1151 124. Wade BS, Pearson PN, Berggren WA, Pälike H. Review and revision of Cenozoic tropical
1152 planktonic foraminiferal biostratigraphy and calibration to the geomagnetic polarity and astronomical
1153 time scale. Earth-Sci Rev. 2011;104(1-3):111-142.
1154 125. Gold DP, Fenton JP, Casas-Gallego M, Novak V, Pérez-Rodríguez I, Cetean C, et al. The
1155 biostratigraphic record of Cretaceous to Paleogene tectono-eustatic relative sea-level change in Jamaica.
1156 J S Am Earth Sci. 2018;86:140-161.
1157 126. Smith SA, Brown JW, Walker JF. So many genes, so little time: A practical approach to
1158 divergence-time estimation in the genomic era. PLoS one. 2018;13(5):e0197433.
1159 127. Dylus DV, Czarkwiani A, Stångberg J, Ortega-Martinez O, Dupont S, Oliveri P. Large-scale gene
1160 expression study in the ophiuroid *Amphiura filiformis* provides insights into evolution of gene regulatory
1161 networks. Evodevo. 2016;7(1):1-17.
1162 128. Reich A, Dunn C, Akasaka K, Wessel G. Phylogenomic analyses of Echinodermata support the
1163 sister groups of Asterozoa and Echinozoa. PLoS One. 2015;10(3):e0119627.
1164 129. Linchangco Jr GV, Foltz DW, Reid R, Williams J, Nodzak C, Kerr AM, et al. The phylogeny of extant
1165 starfish (Asteroidea: Echinodermata) including *Xyloplax*, based on comparative transcriptomics. Mol
1166 Phylogenet Evol. 2017;115:161-170.
1167 130. Janies DA, Witter Z, Linchangco GV, Foltz DW, Miller AK, Kerr AM, et al. EchinoDB, an application
1168 for comparative transcriptomics of deeply-sampled clades of echinoderms. BMC Bioinformatics.
1169 2016;17(1):17-48.
1170 131. Clouse RM, Linchangco Jr GV, Kerr AM, Reid RW, Janies DA. Phylotranscriptomic analysis
1171 uncovers a wealth of tissue inhibitor of metalloproteinases variants in echinoderms. R S Open sci.
1172 2015;2(12):150377.

1173 132. Koga H, Fujitani H, Morino Y, Miyamoto N, Tsuchimoto J, Shibata TF, et al. Experimental
1174 approach reveals the role of *alx1* in the evolution of the echinoderm larval skeleton. PLoS One.
1175 2016;11(2):e0149067.
1176 133. Simakov O, Kawashima T, Marlétaz F, Jenkins J, Koyanagi R, Mitros T, et al. Hemichordate
1177 genomes and deuterostome origins. Nature. 2015;527(7579):459-465.
1178 134. Collins M, Peck LS, Clark MS. Large within, and between, species differences in marine cellular
1179 responses: Unpredictability in a changing environment. Sci Total Environ. 2021:148594.