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]).

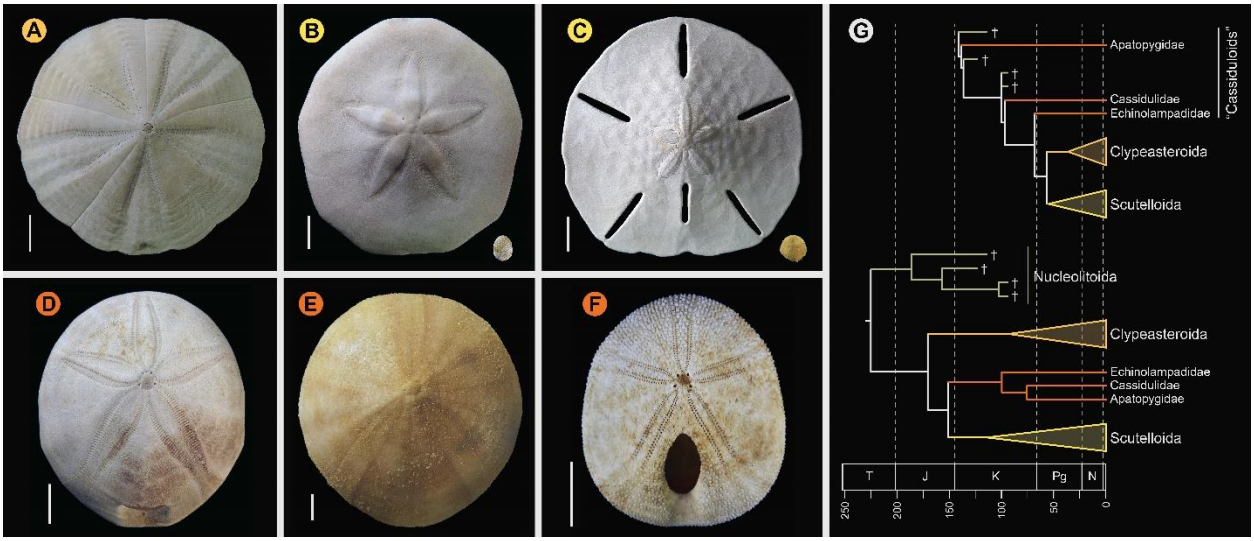

Figure 1: Neognathostomate diversity and phylogenetic relationships. **A.** *Fellaster zelandiae*, North Island, New Zealand (Clypeasteroida). **B.** Large specimen: *Peronella japonica*, Ryukyu Islands, Japan; Small specimen: *Echinocyamus crispus*, Maricaban Island, Philippines (Laganina: Scutelloida). **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 ### (a) Phylogeny of Echinoidea

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

147

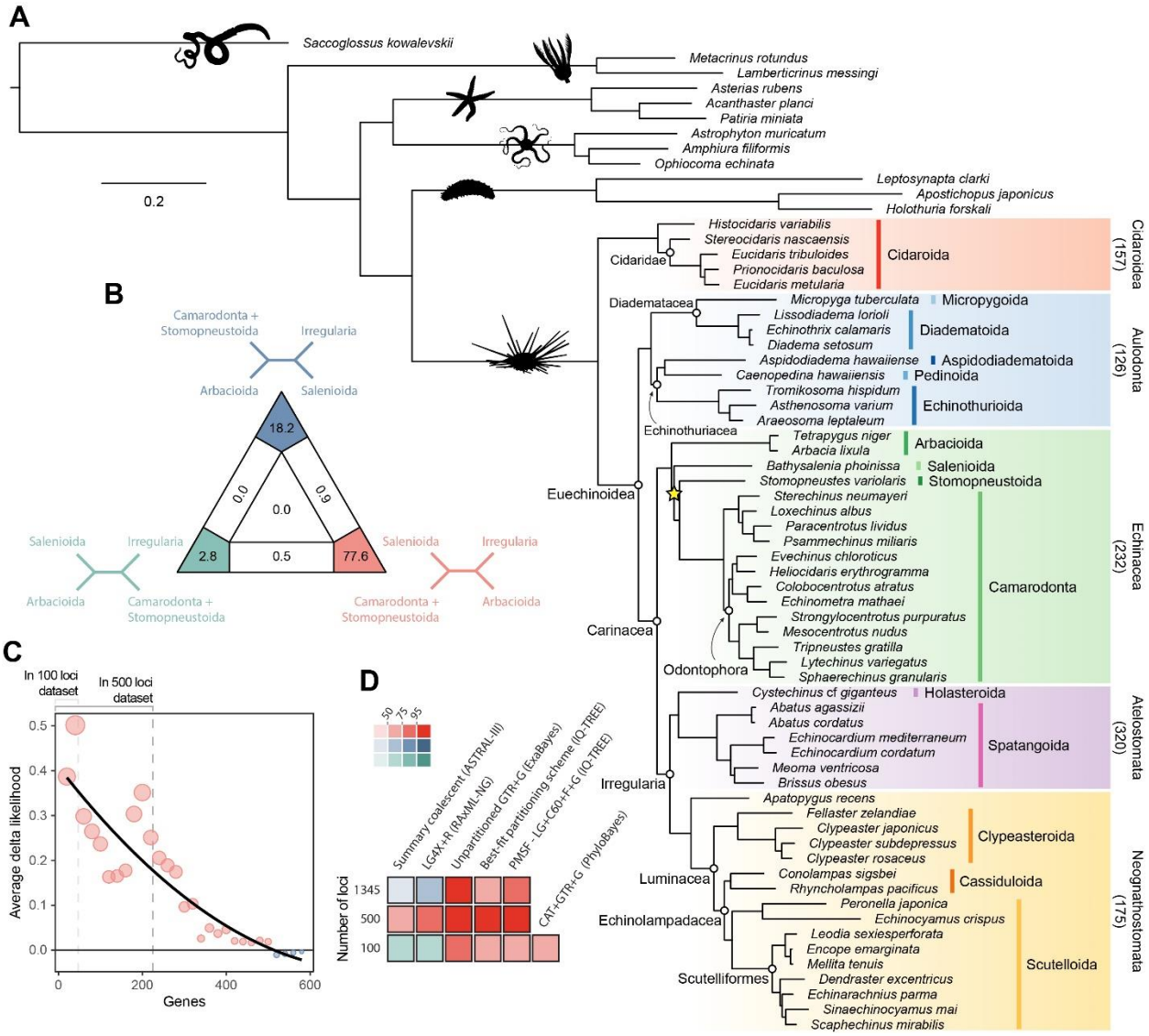

Figure 2: Phylogenetic relationships among major clades of Echinoidea. **A.** Favored topology. With the exception of a single contentious node within Echinacea (marked with a star), all methods supported the same pattern of relationships. Numbers below major clades correspond to the current numbers 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

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

162

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

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

184

185 **(b) Sensitivity of node ages**

186 While alternative methods of inference had minor effects on phylogenetic relationships, they
187 did impact the reconstruction of branch lengths (Fig. S1). Site-heterogeneous models (such as
188 CAT+GTR+G) returned longer branch lengths overall, but also uncovered a larger degree of molecular
189 change among echinoderm classes. Branches connecting these clades were stretched to a much larger
190 extent than those within the ingroup (Fig. S1), which might affect the inference of node dates. We
191 tested this hypothesis by exploring the sensitivity of divergence times to the use of alternative models of
192 molecular evolution (site-homogeneous vs. site-heterogeneous), as well as different clock priors
193 (autocorrelated vs. uncorrelated), and gene sampling strategies (using five different approaches; see
194 Methods). All combinations of these factors were explored, resulting in 20 different time-calibration
195 settings that were run using Bayesian approaches under a constrained tree topology (that shown in Fig.
196 2A). Posterior distributions of node ages were then obtained from each of these time-calibrated
197 analyses, confirming that inferred divergence times were significantly affected by all of these factors
198 (multivariate analysis of variance; all $P < 2.2\text{x}10^{-16}$). While nodes connecting some outgroup taxa were
199 among those most sensitive to these decisions, large effects were also seen among nodes relating to the
200 origin and diversification of the echinoid clades Cidaroidea and Aulodonta. All of these nodes varied in
201 age by more than 60 Ma—and up to 115 Ma—among the consensus topologies of different analyses
202 (Fig. S2).

203

204 In order to isolate and visualize the impact of each of these factors on divergence time
205 estimation, chronograms were represented in a multidimensional space of node dates, with each axis
206 representing the age of a given node. We term this type of graph a chronospace given its similarities to
207 the treespaces commonly used to explore topological differences between phylogenetic trees [39]. Each
208 observation (chronogram) was classified as obtained under a specific clock prior, model of molecular
209 evolution, and gene sampling strategy, and the major effects of each of these choices were extracted
210 with the use of between-group principal component analyses (bgPCAs). This method of ordination
211 allows for the isolation of axes that capture differences between group means, and has been used
212 before with other high-dimensional datasets such as obtained from microarrays and geometric
213 morphometric approaches [40, 41]. The single dimension of chronospace maximizing the distinctiveness
214 of chronograms obtained under different clock models explained 53.7% of the total variance in node
215 ages across all analyses (Fig. 3). In contrast, the choice of different loci or models of molecular evolution
216 showed a much lesser effect, explaining only 10.5% and 4.0% of the total variance, respectively. Even
217 though all of these decisions affected a similar set of sensitive nodes (those mentioned above, as well as
218 some relationships within Atelostomata), the choice of clock model modified the ages of 17 of these by
219 more than 20 Ma (Fig. S3). This degree of change was induced on only 5 nodes by selecting alternative
220 loci (Fig. S4), and was not induced on any node by enforcing different models of evolution (Fig. S5).
221 Regarding gene choice, the ages most different to those obtained under random loci selection were
222 found when using the most clock-like genes (Fig. 3C).

223

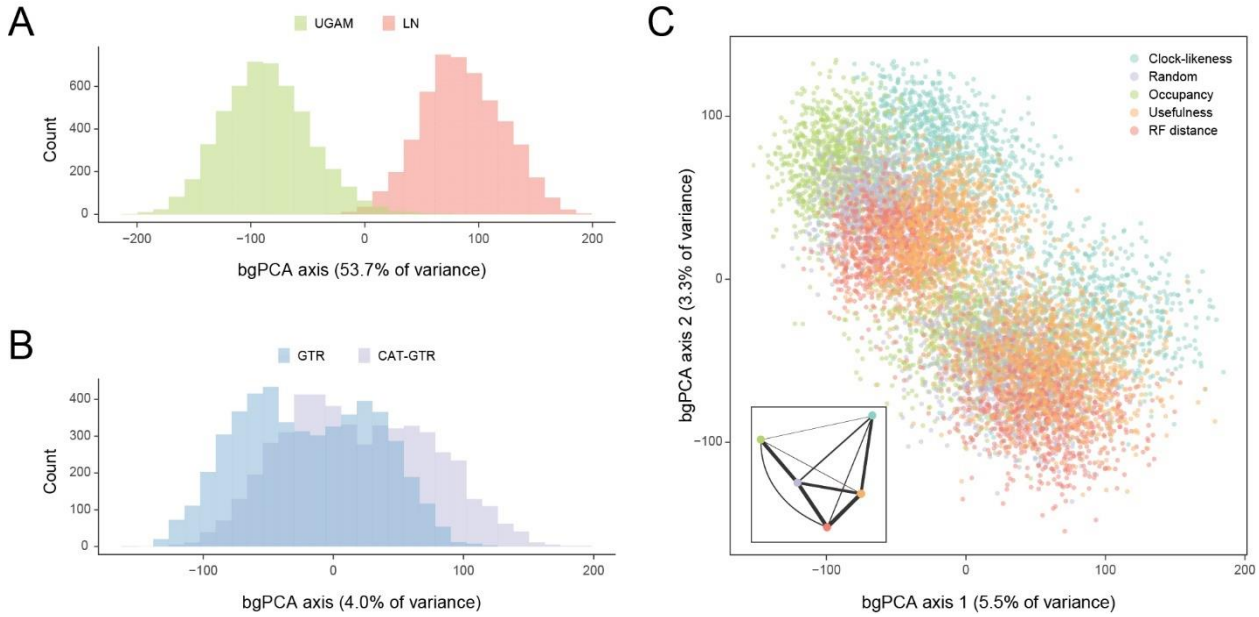

Figure 3: Sensitivity of divergence time estimation to methodological decisions. Between-group principal component analysis (bgPCA) was used to retrieve axes that separate chronograms based on the clock model (**A**), model of molecular evolution (**B**), and gene sampling strategy (**C**) employed. In the latter case, only the first two out of four bgPCA dimensions are shown. The inset shows the centroid for each 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 proportions of total variance explained are shown on the axis labels. The impact of the clock model is such that a bimodal distribution of chronograms can be seen even when bgPCA are built to discriminate based on other factors (as in **C**).

(c) Echinoid (and echinoderm) divergence times

Even when the age of crown Echinodermata was constrained to postdate the appearance of stereom (the characteristic skeletal microstructure of echinoderms) in the Early Cambrian [42, 43], only analyses using the most clock-like loci recovered ages concordant with this (i.e., median ages younger

239 than the calibration enforced; Fig. S4). Instead, most consensus trees favored markedly older ages for
240 the clade, in some cases even predating the origin of the Ediacaran biota [44] (Fig. S6). Despite the
241 relative sensitivity of many of the earliest nodes to methodological choices (Fig. S2), the split between
242 Crinoidea and all other echinoderms (Eleutherozoa) is always inferred to have predated the end of the
243 Cambrian (youngest median age = 492.2 Ma), and the divergence among the other major lineages
244 (classes) of extant echinoderms are constrained to have happened between the Late Cambrian and
245 Middle Ordovician (Fig. S6). Our results also recover an early origin of crown group Holothuroidea (sea
246 cucumbers; range of median ages = 351.6–383.2 Ma), well before the crown groups of other extant
247 echinoderm classes. These dates markedly postdate the first records of holothuroid calcareous rings in
248 the fossil record [45, 46], and imply that this trait does not define the holothuroid crown group but
249 instead evolved from an echinoid-like jaw-apparatus along its stem [47]. The other noteworthy
250 disagreement between our results and those of previous studies [48] involves dating crown group
251 Crinoidea to times that precede the P-T mass extinction (range of median ages = 268.0–327.7 Ma,
252 although highest posterior density intervals are always wide and include Triassic ages).

253 Across all of the analyses performed, the echinoid crown group is found to have originated
254 somewhere between the Pennsylvanian and Cisuralian, with 42% posterior probability falling within the
255 late Carboniferous and 58% within the early Permian (Fig. 4). An origin of the clade postdating the P-T
256 mass extinction is never recovered, even when such ages were allowed under the joint prior (Fig. S7).
257 While the posterior distribution of ages for Euechinoidea spans both sides of the P-T boundary, the
258 remaining earliest splits within the echinoid tree are constrained to have occurred during the Triassic,
259 including the origins of Aulodonta, Carinacea, Echinacea, and Irregularia (Figs. 4 and S6). Many echinoid
260 orders are also inferred to have diverged from their respective sister clades during this period, including
261 aspidodiadematoids, pedinoids, echinothurioids, arbacioids, and salenioids. LTT plots confirm that
262 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).

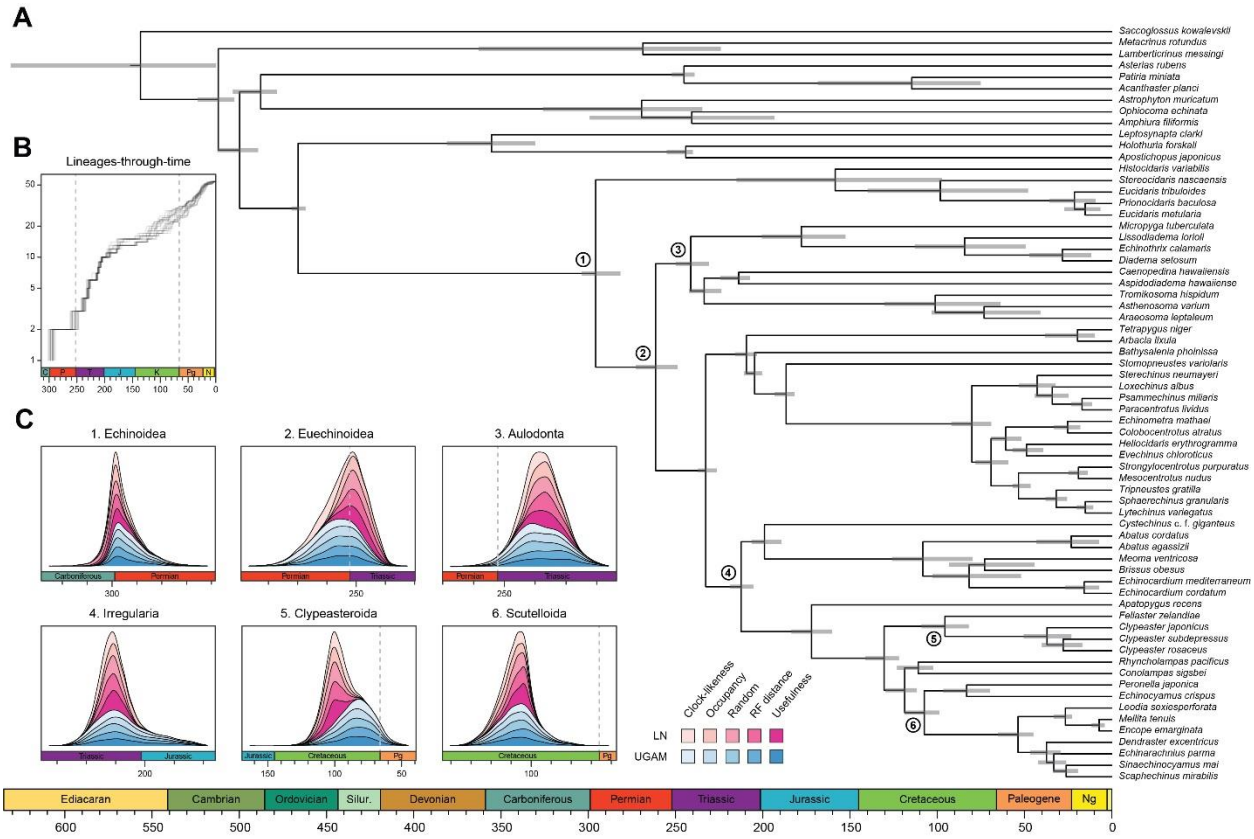

Figure 4: Divergence times among major clades of Echinoidea and other echinoderms. **A.** Consensus chronogram of the two runs using clock-like genes under a CAT+GTR+G model of molecular evolution and a autocorrelated log-normal (LN) clock prior. Node ages correspond to median values, and bars show the 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

275 the K-Pg boundary). Each line corresponds to an individual consensus topology from among the time-
276 calibrated runs performed. **C.** Posterior distributions of the ages of selected nodes (identified in **A**). The
277 effects of models of molecular evolution are not shown, as they represent the least important factor
278 (see Fig. 3); distributions using site-homogeneous and heterogeneous models are combined for every
279 combination of targeted loci and clock prior.

280

281 ## Discussion

282 ### (a) The echinoid Tree of Life

283 In agreement with previous phylogenomic studies [3, 4], echinoid diversity can be subdivided
284 into five major clades (Fig. 2A). Cidaroids form the sister group to all other crown group echinoids
285 (Euechinoidea). Some aspects of the relationships among sampled cidaroids are consistent with previous
286 molecular [49] and morphological studies [19], including an initial split between *Histocidaris* and the
287 remaining taxa, representing the two main branches of extant cidaroids [5, 35]. Others, such as the
288 nested position of *Prionocidaris baculosa* within the genus *Eucidaris,* not only implies paraphyly of this
289 genus but also suggests the need for a taxonomic reorganization of the family Cidaridae. Within
290 euechinoids, the monophyly of Aulodonta is supported for the first time with sampling of all of its major
291 groups. The subdivision of these into a clade that includes diadematoids plus micropygoids (which we
292 propose should retain the name Diadematacea), sister to a clade including echinothurioids and
293 pedinoids (Echinothuriacea *sensu* [4]) is strongly reminiscent of some early classifications (e.g., [50]).
294 Our expanded phylogenomic sampling also confirms an aulodont affinity for aspidodiadematoids [3, 35]
295 and places them within Echinothuriacea as the sister group to Pedinoida.

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

309 The remaining diversity of echinoids, which forms the clade Carinacea (Fig. 2), is subdivided into
310 Irregularia and their sister clade among regulars, for which we amend the name Echinacea to include
311 Salenioida. Given the striking morphological gap separating regular and irregular echinoids, the origin of
312 Irregularia has been shrouded in mystery [19, 23, 50]. Our complete sampling of major regular lineages
313 determines Echinacea *sensu stricto* to be the sister clade to irregular echinoids. A monophyletic
314 Echinacea was also supported in a recent total-evidence analysis [3] , but the incomplete molecular
315 sampling of that study resulted in a slightly different topology that placed salenioids as the sister group
316 to the remaining lineages. However, an overall lack of morphological synapomorphies uniting these
317 clades had previously been acknowledged [19]. While the relationships within Echinacea proved to be
318 difficult to resolve even with thousands of loci, multiple lines of evidence lead us to prefer a topology in

319 which salenioids form a clade with camarodonts + stomopneustoids, with arbacioids sister to all of these
320 (Fig. 2).

321 As has been already established [3, 4, 14, 19, 30], the lineages of irregular echinoids here
322 sampled are subdivided into Atelostomata (heart urchins and allies) and Neognathostomata (sand
323 dollars, sea biscuits, and "cassiduloids"). Despite the former being the most diverse of the five main
324 clades of echinoids (Fig. 2), its representation in phylogenomic studies remains low, and its internal
325 phylogeny poorly constrained [35]. On the contrary, recent molecular studies have greatly improved our
326 understanding of the relationships among neognathosomates [3, 4, 36], revealing an evolutionary
327 history that dramatically departs from previous conceptions. Even when scutelloids and clypeasteroids
328 were never recovered as reciprocal sister lineages by molecular phylogenies (e.g., [13, 14, 30]), this
329 result was not fully accepted until phylogenomic data confidently placed echinolampadids as the sister
330 lineage to sand dollars [4]. At the same time, this result rendered the position of the remaining
331 "cassiduloids", a taxonomic wastebasket with an already complicated history of classification [19, 27, 28,
332 51], entirely uncertain. An attempt to constrain the position of these using a total-evidence approach [3]
333 subdivided the "cassiduloids" into three unrelated clades: Nucleolitoida, composed of extinct lineages
334 and placed outside the node defined by Scutelloida + Clypeasteroida, and two other clades nested
335 within it (see Fig. 1). Extant "cassiduloids" were recovered as members of one of the latter clades,
336 representing the monophyletic sister group to sand dollars. Here, we show that *Apatopygus recens* does
337 not belong within this clade but is instead the sister group to all other extant neognathostomates. Given
338 this phylogenetic position, as well as the morphological similarities between *Apatopygus* and the
339 entirely extinct nucleolitids [19, 28, 51-53], it is likely that the three extant species of apatopygids
340 represent the last surviving remnants of Nucleolitoida, a clade of otherwise predominantly Mesozoic
341 neognathostomates [3]. Because of the renewed importance in recognizing this topology, we propose
342 the name Luminacea for the clade uniting all extant neognathostomates with the exclusion of

343 *Apatopygus* (Fig. 2A). This nomenclature refers to the dynamic evolutionary history of the Aristotle's
344 lantern (i.e., the echinoid jaw-apparatus) within the clade (present in the adults of both clypeasteroids
345 and scutelloids, but found only in the juveniles of Cassiduloida *sensu stricto*), the inclusion of the so-
346 called lamp urchins (echinolampadids) within the clade, and the illumination provided by this hitherto
347 unexpected topology. The previous misplacement of *Apatopygus* within this clade ([3]; see Fig. 1G) is
348 likely a consequence of tip-dating preferring more stratigraphically congruent topologies [54], an effect
349 that can incorrectly resolve taxa on long terminal branches [55]. Given the generally useful phylogenetic
350 signal of stratigraphic information [56], this inaccuracy further highlights the unusual evolutionary
351 history of living apatopygids.

352

353 **(b) Chronospaces: a statistical exploration of time-calibration strategies**

354 Calibrating phylogenies to absolute time is crucial to understanding evolutionary history, as the
355 resulting chronograms provide a major avenue for testing hypotheses of diversification, character
356 evolution and other macroevolutionary processes. However, the accuracy and precision of the inferred
357 divergence times hinge upon many methodological choices that are often difficult or time-consuming to
358 justify (e.g., calibration strategies, prior distributions on node ages, clock models, etc.) [57-62].
359 Understanding the impact of each of these sources of uncertainty can be difficult. The sensitivity of
360 inferred ages is commonly explored by running analyses under different settings and summarizing the
361 results in tables or by stacking chronograms in order to visualize the relative position of nodes (see for
362 example [58, 59]). This approach discards a lot of information contained in Bayesian posterior
363 distributions of time-calibrated topologies. For example, patterns of correlation among branch lengths
364 (i.e., ways in which changes in a given branch length are accommodated by other regions of the tree)
365 are lost, and alternative ways in which node ages are attained (e.g., which of all descendant branches

366 expand as a given node becomes older) are not usually explored, even though they can potentially
367 represent very different evolutionary scenarios.

368 Here we analyze the sensitivity of node ages to alternative criteria to sample loci from
369 phylogenomic datasets, as well as different ways of modelling the patterns of molecular evolution
370 across sites and the rates of evolution among lineages. To do so, we introduce an approach to visualize
371 the distribution of chronograms in a multidimensional space of node ages, a chronospace, and measure
372 the overall effect of these decisions on inferred dates using multivariate statistical methods. Our results
373 reveal that divergence times obtained under site-homogeneous and site-heterogeneous models (such as
374 CAT+GTR+G) are broadly comparable. This happens despite the latter estimating higher levels of
375 sequence divergence and stretching branches in a non-isometric manner (Fig. S1). Site-heterogeneous
376 models have become common for the analysis of phylogenomic datasets spanning deep divergences, as
377 they exhibit reduced sensitivity to saturation and are thus more robust to long-branch attraction
378 artifacts [63-65]. These models are also central to the debate regarding contentious nodes in the tree of
379 life, including the position of ctenophores [66, 67] and the monophyly of arachnids [68, 69]. While the
380 use of CAT models for time calibration has become equally popular, the degree to which they impact
381 estimates of node ages has received less scrutiny. The lack of a meaningful effect uncovered here,
382 coupled with their high computational burden [70], questions their usefulness for time-scaling
383 phylogenies. A similar result was recently found when comparing site-homogeneous models with
384 different numbers of parameters [71], suggesting that relaxed clocks adjust branch rates in a manner
385 that buffers the effects introduced by using more complex (and computationally-intensive) models of
386 sequence evolution.

387 Similarly, the choice between different loci has a small effect on inferred ages, with little
388 evidence of a systematic difference between the divergence times supported by randomly-chosen loci
389 and those found using targeted sampling criteria, such as selecting genes for their phylogenetic signal,

usefulness, occupancy, or clock-likeness. A meaningful effect was restricted to a few ancient nodes (e.g., Echinodermata), for which clock-like genes suggested younger ages that are more consistent with fossil evidence. This result provides direct empirical evidence that validates the selection of clock-like genes for inferring deep histories of diversification, an approach so far supported mostly by simulations [62, 72]. At the same time, the choice of loci seems to be less relevant for younger nodes, at least when employing large numbers of characters as done here. Finally, the choice between alternative clock models induced 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., [73]) or—as done here—focusing on results that are robust to them.

(c) Echinoid origins and diversification

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 [74] 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.

412 The late Paleozoic and Triassic origins inferred for the crown group and many euechinoid orders
413 prompts a re-evaluation of fossils from this interval of time. Incompletely known fossil taxa such as the
414 Pennsylvanian *?Eotiaris meurevillensis*, with an overall morphology akin to that of crown group
415 echinoids, has a stratigraphic range consistent with our inferred date for the origin of the echinoid
416 crown group [20]. Additionally, the Triassic fossil record of echinoids has been considered to be
417 dominated by stem group cidaroids, with the first euechinoids not known until the Late Triassic [17, 75].
418 However, the Triassic origins of many euechinoid lineages supported by our analyses necessitates that
419 potential euechinoid affinities should be re-considered for this diversity of Triassic fossils. This is
420 especially the case for the serpianotiarids and triadocidarids, abundant Triassic families variously
421 interpreted as cidaroids, euechinoids, or even stem echinoids [3, 17, 75, 76]. A reinterpretation of any of
422 these as euechinoids would suggest that the long-implied gap in the euechinoid record [16, 18] is caused
423 by our inability to correctly place these key fossils, as opposed to an incompleteness of the fossil record
424 itself.

425 While our phylogenomic approach is the first to resolve the position of all major cassiduloid
426 lineages, the inferred ages for many nodes within Neognathostomata remain in strong disagreement
427 with the fossil record. No Mesozoic fossil can be unambiguously assigned to either sand dollars or sea
428 biscuits, a surprising situation given the good fossilization potential and highly distinctive morphology of
429 these clades [19, 21, 25, 77]. While molecular support for a sister group relationship between scutelloids
430 and echinolampadids already implied this clade (Echinolampadacea) must have split from clypeasteroids
431 by the Late Cretaceous [3, 4, 14, 19], this still left open the possibility that the crown groups of sand
432 dollars and sea biscuits radiated in the Cenozoic. Under this scenario, the Mesozoic history of these
433 groups could have been comprised of forms lacking their distinctive morphological features,
434 complicating their correct identification. This hypothesis is here rejected, with the data unambiguously
435 supporting the radiation of both crown groups preceding the K-Pg mass extinction (Fig. 4C). While it

436 remains possible that these results are incorrect even after such a thorough exploration of the time-
437 calibration toolkit (see for example [62, 78]), these findings call for a critical reassessment of the
438 Cretaceous fossil record, and a better understanding of the timing and pattern of morphological
439 evolution among fossil and extant neognathostomates. For example, isolated teeth with an overall
440 resemblance to those of modern sand dollars and sea biscuits have been found in Lower Cretaceous
441 deposits [79], raising the possibility that other overlooked and disarticulated remains might close the
442 gap between rocks and clocks.

443

444 **(d) Conclusions**

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

454 Although divergence time estimation is known to be sensitive to many methodological
455 decisions, systematically quantifying the relative impact of these on inferred ages has rarely been done.
456 Here we propose an approach based on chronospaces (implemented in R [80] and available at
457 https://github.com/mongiardino/chronospace) that can help visualize key effects and determine the
458 sensitivity of node dates to different assumptions. Our results shed new light on the early evolutionary

459 history of crown group echinoids and its relationship with the P-T mass extinction event, a point in time
460 where the fossil record provides ambiguous answers. They also establish with confidence a Cretaceous
461 origin for the sand dollars and sea biscuits, preceding their first appearance in the fossil record by at
462 least 40 to 50 Ma, respectively (and potentially up to 65 Ma). These clades, therefore, join several well-
463 established cases of discrepancies between the fossil record and molecular clocks, such as those
464 underlying the origins of placental mammals [81] and flowering plants [82].

465

466 # Methods

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

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

475 Raw reads for all transcriptomic datasets were trimmed or excluded using quality scores with
476 Trimmomatic v. 0.36 [83] under default parameters. Further sanitation steps were performed using the
477 Agalma 2.0 phylogenomic workflow [84], and datasets were assembled *de novo* with Trinity 2.5.1 [85].
478 For genomic shotgun sequences, adapters were removed with BBDuk
479 (https://sourceforge.net/projects/bbmap/), and UrQt v. 1.0.18 [86] was used to filter short reads (size <
480 50) and trim low-quality ends (score < 28). Datasets were then assembled using MEGAHIT v. 1.1.2 [87].

481 Draft genomes were masked using RepeatMasker v. 4.1.0 [88, 89], before obtaining gene predictions
482 with AUGUSTUS 3.2.3 [90]. A custom set of universal single-copy orthologs (USCOs) obtained from the
483 latest *Strongylocentrotus purpuratus* genome assembly (Spur v. 5.0) was employed as the training
484 dataset. Settings and further details of these analyses can be found in the Supplementary Information.

485 Multiplexed transcriptomes were sanitized from cross-contaminants using CroCo v. 1.1 [91], and
486 likely non-metazoan contaminants were removed using alien_index v. 3.0 [92] (removing sequences
487 with AI values > 45). Datasets were imported back into Agalma, which automated orthology inference
488 (as described in [84, 93]), gene alignment with MACSE [94], and trimming with GBLOCKS [95]. The amino
489 acid supermatrix was reduced using a 70% occupancy threshold, producing a final dataset of 1,346 loci
490 (327,695 sites). As a final sanitation step, gene trees were obtained using ParGenes v. 1.0.1 [96], which
491 performed model selection (minimizing the Bayesian Information Criterion) and inference using 100
492 bootstrap (BS) replicates. Trees were then used to remove outlier sequences with TreeShrink v. 1.3.1
493 [97]. We specified a reduced tolerance for false positives and limited removal to at most three terminals
494 which had to increase tree diameter by at least 25% (-q 0.01 -k 3 -b 25). Statistics for the supermatrix
495 and all assemblies can be found in Table S2.

496

497 **(b) Phylogenetic inference**

498 Inference was performed under multiple probabilistic and coalescent-aware methods, known to
499 differ in their susceptibility to model violations. Coalescent-based inference was performed using the
500 summary method ASTRAL-III [98], estimating support as local posterior probabilities [99]. Among
501 concatenation approaches, we used Bayesian inference under an unpartitioned GTR+G model in
502 ExaBayes 1.5 [100]. Two chains were run for 2.5 million generations, samples were drawn every one
503 hundred, and the initial 25% was discarded as burn-in. We also explored maximum likelihood inference

504 with partitioned and unpartitioned models. For the former, the fast-relaxed clustering algorithm was
505 used to find the best-fitting model among the top 10% using IQ-TREE 1.6.12 [101, 102] (-m MFP+MERGE
506 -rclusterf 10 -rcluster-max 3000), and support was evaluated with 1,000 ultrafast BS replicates [103]. For
507 the latter, we used the LG4X+R model in RAxML-NG v. 0.5.1 [104] and evaluated support with 200
508 replicates of BS. Finally, we also implemented the site-heterogeneous LG+C60+F+G mixture model using
509 the posterior mean site frequency (PMSF) approach to provide a fast approximation of the full profile
510 mixture model [105], allowing the use of 100 BS replicates to estimate support. Given some degree of
511 topological conflict between the results of the other methods (see below), multiple guide trees were
512 used to estimate site frequency profiles, but the resulting phylogenies were identical.

513 Given conflicts between methods in the resolution of one particular node (involving the
514 relationships among Arbacioida, Salenioida, and the clade of Stomopneustoida + Camarodonta), all
515 methods were repeated after reducing the matrix to 500 and 100 loci selected for their phylogenetic
516 usefulness using the approach described in [3, 37] and implemented in the *genesortR* script
517 (https://github.com/mongiardino/genesortR). This approach relies on seven gene properties routinely
518 used for phylogenomic subsampling, including multiple proxies for phylogenetic signal—such as the
519 average BS and Robinson-Foulds (RF) similarity to a target topology—as well as several potential sources
520 of systematic bias (e.g., rate and compositional heterogeneities). Outgroups were removed before
521 calculating these metrics. RF similarity was measured to a species tree that had the conflicting
522 relationship collapsed so as not to bias gene selection in favor of any resolution. A principal component
523 analysis (PCA) of this dataset resulted in a dimension (PC 2, 17.6% of variance) along which phylogenetic
524 signal increased while sources of bias decreased (Fig. S9), and which was used for loci selection. For the
525 smallest dataset, we also performed inference under the site-heterogeneous CAT+GTR+G model using
526 PhyloBayes-MPI [106]. Three runs were continued for > 10,000 generations, sampling every two

527 generations and discarding the initial 25%. Convergence was confirmed given a maximum bipartition
528 discrepancy of 0.067 and effective sample sizes for all parameters > 150.

529 Two other approaches were implemented in order to assist in resolving the contentious node.
530 First, we implemented a likelihood-mapping analysis [107] in IQ-TREE to visualize the phylogenetic signal
531 for alternative resolutions of the quartet involving these three lineages (Arbacioida, Salenioida, and
532 Stomopneustoida + Camarodonta) and their sister clade (Irregularia; other taxa were excluded). Second,
533 we estimated the log-likelihood scores of each site in RAxML (using best-fitting models) for the two most
534 strongly supported resolutions found through likelihood mapping. These were used to calculate gene-
535 wise differences in scores, or δ values [108]. In order to search for discernable trends in the signal for
536 alternative topologies, genes were ordered based on their phylogenetic usefulness (see above) and the
537 mean per-locus δ values of datasets composed of multiples of 20 loci (i.e., the most useful 20, 40, etc.)
538 were calculated.

539

540 **(c) Time calibration**

541 Node dating was performed using relaxed molecular clocks in PhyloBayes v4.1 using a fixed
542 topology and a novel set of 22 fossil calibrations corresponding to nodes from our newly inferred
543 phylogeny (listed in the Supplementary Information). Depending on the node, we enforced both
544 minimum and maximum bounds, or either one of these. A birth-death prior was used for divergence
545 times, which allowed for the implementation of soft bounds [109], leaving 5% prior probability of
546 divergences falling outside of the specified interval. We explored the sensitivity of divergence time
547 estimates to gene selection, model of molecular evolution, and clock prior. One hundred loci were
548 sampled from the full supermatrix according to four targeted sampling schemes: usefulness (calculated
549 as explained above, except incorporating all echinoderm terminals), phylogenetic signal (i.e., smallest RF

550 distance to species tree), clock-likeness (i.e., smallest variance of root-to-tip distances), and level of
551 occupancy. For clock-likeness, we only considered loci that lay within one standard deviation of the
552 mean rate (estimated by dividing total tree length by the number of terminals [110]), as this method is
553 otherwise prone to selecting largely uninformative loci (Fig. S10; [37]). A fifth sample of randomly
554 chosen loci was also evaluated. These five datasets were run under two models of molecular evolution,
555 the site-homogeneous GTR+G and the site-heterogeneous CAT+GTR+G, and both uncorrelated gamma
556 (UGAM) and autocorrelated log-normal (LN) clocks were implemented. The combination of these
557 settings (loci sampled, model of evolution, and clock prior) resulted in 20 analyses. For each, two runs
558 were continued for 20,000 generations, after which the initial 25% was discarded and the chains thinned
559 to every two generations (see log-likelihood trace plots in the electronic supplementary material, Fig.
560 S11). To explore the sensitivity of divergence times to these methodological decisions, 500 random
561 chronograms were sampled from each analysis (250 from each run), and their node dates were
562 subjected to multivariate analyses of variance (MANOVA) and between-group PCA (bgPCA) using
563 package *Morpho* [111] in the R statistical environment [80]. bgPCA involves the use of PCA on the
564 covariance matrix of group means (e.g., the mean ages of all nodes obtained under UGAM or LN clocks),
565 followed by the projection of original samples onto the obtained bgPC axes. The result is a
566 multidimensional representation of divergence times—a chronospace—rotated so as to capture the
567 distinctiveness of observations obtained under different settings. Separate bgPCAs were performed for
568 loci sampling strategy, clock prior and model of molecular evolution, and the proportion of total
569 variance explained by bgPC axes was interpreted as an estimate of the relative impact of these choices
570 on inferred times of divergence. Finally, lineage-through-time plots were generated using *ape* [112].

571

572 **Author contributions**

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

Data, code and materials

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

Competing interests

We declare no competing interests.

Acknowledgments

We are grateful to the crew of R/Vs Antéa, Atlantis, Falkor and Mirai, HOV Alvin, and the crew and PIs of BIOMAGLO deep-sea cruises (Laure Corbari, Karine Olu-Le Roy, Sarah Samadi) for assistance in specimen collection. We would also like to thank Libby Liggins, Wilma Blom, Owen Anderson, Francisco Solís-Marín, Carlos A. Conejeros-Vargas, and Jih-Pai Lin for the collection and shipment of specimens. We gratefully acknowledge assistance from Thomas Saucède (Univ. Burgundy), Marc Eléaume (MNHN),

and Charlotte Seid (Scripps) in accessing, cataloging and processing museum specimens. Tim Ravasi provided resources and collection facilities to GWR via King Abdullah University of Science and Technology (KAUST). Institutional support was provided by the Central Research Laboratories and the Department of Geology and Palaeontology at the Natural History Museum Vienna, Austria. This work was supported by a Yale Institute for Biospheric Studies Doctoral Dissertation Improvement Grant and a Society of Systematic Biologists Graduate Student Research Award to NMK. AK received funding from an Austrian Science Fund project (FWF, project number P 29508-B25), FA from Agencia Nacional de Investigación (ANID/PAI Inserción en la Academia, project number PAI79170033), and GWR and RM from NSF grant DEB-2036186. JRT was supported by a Royal Society Newton International Fellowship and a Leverhulme Trust Early Career Fellowship. NMK was supported by a Yale University fellowship.

References

1. Donoghue PC, Benton MJ (2007) **Rocks and clocks: calibrating the Tree of Life using fossils and molecules** *Trends in Ecology & Evolution* 22:424-431.

2. Smith AB, Peterson KJ (2002) **Dating the time of origin of major clades: molecular clocks and the fossil record** *Annual Review of Earth and Planetary Sciences* 30:65-88.

3. Mongiardino Koch N, Thompson JR (2021) **A total-evidence dated phylogeny of Echinoidea combining phylogenomic and paleontological data** *Systematic Biology* 70:421-439.

4. Mongiardino Koch N, Coppard SE, Lessios HA, Briggs DE, Mooi R, Rouse GW (2018) **A phylogenomic resolution of the sea urchin tree of life** *BMC Evolutionary Biology* 18:189.

5. Kroh A, Mooi R (2021) **World Echinoidea Database**. Visited 13/July/2021. http://www.marinespecies.org/echinoidea.

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

617 7. Edmunds PJ, Carpenter RC (2001) **Recovery of *Diadema antillarum* reduces macroalgal cover**
618 **and increases abundance of juvenile corals on a Caribbean reef** *Proceedings of the National Academy*
619 *of Sciences* 98:5067-5071.
620 8. Harrold C, Pearse JS (1987) **The ecological role of echinoderms in kelp forests** In: Jangoux M,
621 Lawrence JM, editors. *Echinoderm Studies 2*. Rotterdam: AA Balkema. pp. 137–233.
622 9. Zonneveld JP, Furlong CM, Sanders SC (2016) **Triassic echinoids (Echinodermata) from the**
623 **Aksala Formation, north Lake Laberge, Yukon Territory, Canada** *Papers in Palaeontology* 2:87-100.
624 10. Thompson JR, Petsios E, Bottjer DJ (2017) **A diverse assemblage of Permian echinoids**
625 **(Echinodermata, Echinoidea) and implications for character evolution in early crown group echinoids**
626 *Journal of Paleontology* 91:767-780.
627 11. Kier PM (1977) **Triassic echinoids** *Smithsonian Contributions to Paleobioliology* 30:1-88.
628 12. Twitchett RJ, Oji T (2005) **Early Triassic recovery of echinoderms** *Comptes Rendus Palevol* 4:531-
629 542.
630 13. Thompson JR, Erkenbrack EM, Hinman VF, McCauley BS, Petsios E, Bottjer DJ (2017)
631 **Paleogenomics of echinoids reveals an ancient origin for the double-negative specification of**
632 **micromeres in sea urchins** *Proceedings of the National Academy of Sciences* 114:5870-5877.
633 14. Smith AB, Pisani D, Mackenzie-Dodds JA, Stockley B, Webster BL, Littlewood DTJ (2006) **Testing**
634 **the molecular clock: molecular and paleontological estimates of divergence times in the Echinoidea**
635 **(Echinodermata)** *Molecular Biology and Evolution* 23:1832-1851.
636 15. Smith A, Hollingworth N (1990) **Tooth structure and phylogeny of the Upper Permian echinoid**
637 ***Miocidaris keyserlingi*** *Proceedings of the Yorkshire Geological Society* 48:47-60.
638 16. Thompson JR, Petsios E, Davidson EH, Erkenbrack EM, Gao F, Bottjer DJ (2015) **Reorganization**
639 **of sea urchin gene regulatory networks at least 268 million years ago as revealed by oldest fossil**
640 **cidaroid echinoid** *Scientific Reports* 5:15541.

641 17. Smith AB (2007) **Intrinsic versus extrinsic biases in the fossil record: contrasting the fossil**
642 **record of echinoids in the Triassic and early Jurassic using sampling data, phylogenetic analysis, and**
643 **molecular clocks** *Paleobiology* 33:310-323.
644 18. Thompson JR, Hu S-x, Zhang Q-Y, Petsios E, Cotton LJ, Huang J-Y, et al (2018) **A new stem group**
645 **echinoid from the Triassic of China leads to a revised macroevolutionary history of echinoids during**
646 **the end-Permian mass extinction** *Royal Society Open Science* 5:171548.
647 19. Kroh A, Smith AB (2010) **The phylogeny and classification of post-Palaeozoic echinoids** *Journal*
648 *of Systematic Palaeontology* 8:147-212.
649 20. Thompson JR, Mirantsev GV, Petsios E, Bottjer DJ (2020) **Phylogenetic analysis of the**
650 **Archaeocidaridae and Palaeozoic Miocidaridae (Echinodermata, Echinoidea) and the origin of crown**
651 **group echinoids** *Papers in Palaeontology* 6:217-249.
652 21. Kier PM (1982) **Rapid evolution in echinoids** *Palaeontology* 25:1-9.
653 22. Barras CG (2008) **Morphological innovation associated with the expansion of atelostomate**
654 **irregular echinoids into fine-grained sediments during the Jurassic** *Palaeogeography,*
655 *Palaeoclimatology, Palaeoecology* 263:44-57.
656 23. Saucède T, Mooi R, David B (2007) **Phylogeny and origin of Jurassic irregular echinoids**
657 **(Echinodermata: Echinoidea)** *Geological Magazine* 144:333-359.
658 24. Hopkins MJ, Smith AB (2015) **Dynamic evolutionary change in post-Paleozoic echinoids and the**
659 **importance of scale when interpreting changes in rates of evolution** *Procedings of the National*
660 *Academy of Sciences* 112:3758-3763.
661 25. Mooi R (1990) **Paedomorphosis, Aristotle's lantern, and the origin of the sand dollars**
662 **(Echinodermata: Clypeasteroida)** *Paleobiology* 16:25-48.
663 26. Nebelsick JH (2020) **Ecology of clypeasteroids**. In: Lawrence JM, editor. *Sea Urchins: Biology and*
664 *Ecology, 4th edition*. Cambridge: Academic Press. pp. 315-331.

665 27. Smith AB (2001) **Probing the cassiduloid origins of clypeasteroid echinoids using**
666 **stratigraphically restricted parsimony analysis** *Paleobiology* 27:392-404.
667 28. Souto C, Mooi R, Martins L, Menegola C, Marshall CR (2019) **Homoplasy and extinction: the**
668 **phylogeny of cassidulid echinoids (Echinodermata)** *Zoological Journal of the Linnean Society* 187:622-
669 660.
670 29. Kier PM (1962) **Revision of the cassiduloid echinoids** *Smithsonian Miscellaneous Collections*
671 144:1-262.
672 30. Littlewood D, Smith A (1995) **A combined morphological and molecular phylogeny for sea**
673 **urchins (Echinoidea: Echinodermata)** *Philosophical Transactions of the Royal Society B: Biological*
674 *Sciences* 347:213-234.
675 31. Mongiardino Koch N (2021) **Exploring adaptive landscapes across deep time: A case study**
676 **using echinoid body size** *Evolution* 75:1567-1581.
677 32. Mooi R (1990) **Progenetic miniaturization in the sand dollar *Sinaechinocyamus*: implications**
678 **for clypeasteroid phylogeny.** In: De Ridder C, Dubois P, Lahaye M, Jangoux M, editors. *Echinoderm*
679 *Research*. Rotterdam: AA Balkema. pp. 137-143.
680 33. Dunn CW, Zapata F, Munro C, Siebert S, Hejnol A (2018) **Pairwise comparisons across species**
681 **are problematic when analyzing functional genomic data** *Proceedings of the National Academy of*
682 *Sciences* 115:409-417.
683 34. Smith SD, Pennell MW, Dunn CW, Edwards SV (2020) **Phylogenetics is the new genetics (for**
684 **most of biodiversity)** *Trends in Ecology & Evolution* 35:415-425.
685 35. Kroh A (2020) **Phylogeny and classification of echinoids**. In: Lawrence J, editor. *Sea Urchins:*
686 *Biology and Ecology, 4th edition*. Cambridge: Academic Press. pp. 1-17.

687 36. Lin J-P, Tsai M-H, Kroh A, Trautman A, Machado DJ, Chang L-Y, et al (2020) **The first complete**
688 **mitochondrial genome of the sand dollar Sinaechinocyamus mai (Echinoidea: Clypeasteroida)**
689 *Genomics* 112:1686-1693.
690 37. Mongiardino Koch N (2021) **Phylogenomic subsampling and the search for phylogenetically**
691 **reliable loci** *Molecular Biology and Evolution* msab151. doi: 10.1093/molbev/msab151.
692 38. Dell'Ampio E, Meusemann K, Szucsich NU, Peters RS, Meyer B, Borner J, et al (2014) **Decisive**
693 **data sets in phylogenomics: lessons from studies on the phylogenetic relationships of primarily**
694 **wingless insects** *Molecular Biology and Evolution* 31:239-249.
695 39. Hillis DM, Heath TA, John KS (2005) **Analysis and visualization of tree space** *Systematic Biology*
696 54:471-482.
697 40. Culhane AC, Perriere G, Considine EC, Cotter TG, Higgins DG (2002) **Between-group analysis of**
698 **microarray data** *Bioinformatics* 18:1600-1608.
699 41. Cardini A, O'Higgins P, Rohlf FJ (2019) **Seeing distinct groups where there are none: spurious**
700 **patterns from between-group PCA** *Evolutionary Biology* 46:303-316.
701 42. Zamora S, Lefebvre B, Álvaro JJ, Clausen S, Elicki O, Fatka O, et al (2013) **Cambrian echinoderm**
702 **diversity and palaeobiogeography** *Geological Society, London, Memoirs*. 38:157-171.
703 43. Bottjer DJ, Davidson EH, Peterson KJ, Cameron RA (2006) **Paleogenomics of echinoderms**
704 *Science* 314:956-960.
705 44. Pu JP, Bowring SA, Ramezani J, Myrow P, Raub TD, Landing E, et al (2016) **Dodging snowballs:**
706 **Geochronology of the Gaskiers glaciation and the first appearance of the Ediacaran biota** *Geology*
707 44:955-958.
708 45. Reich M (2015) **Different pathways in early evolution of the holothurian calcareous ring**. In:
709 Zamora S, Rábano I, editors. *Progress in Echinoderm Palaeobiology*. Madrid: Instituto Geológico y
710 Minero de España. pp. 137-145.

711 46. Miller AK, Kerr AM, Paulay G, Reich M, Wilson NG, Carvajal JI, et al (2017) **Molecular phylogeny**
712 **of extant Holothuroidea (Echinodermata)** *Molecular Phylogenetics and* Evolution 111:110-131.
713 47. Rahman IA, Thompson JR, Briggs DE, Siveter DJ, Siveter DJ, Sutton MD (2019) **A new**
714 **ophiocistioid with soft-tissue preservation from the Silurian Herefordshire Lagerstätte, and the**
715 **evolution of the holothurian body plan** *Proceedings of the Royal Society* B 286:20182792.
716 48. Rouse GW, Jermiin LS, Wilson NG, Eeckhaut I, Lanterbecq D, Oji T, et al (2013) **Fixed, free, and**
717 **fixed: the fickle phylogeny of extant Crinoidea (Echinodermata) and their Permian–Triassic origin**.
718 *Molecular Phylogenetics and Evolution* 66:161-181.
719 49. Brosseau O, Murienne J, Pichon D, Vidal N, Eléaume M, Ameziane N (2012) **Phylogeny of**
720 **Cidaroida (Echinodermata: Echinoidea) based on mitochondrial and nuclear markers** *Organisms*
721 *Diversity & Evolution* 12:155-165.
722 50. Durham JW, Melville R (1957) **A classification of echinoids** *Journal of Paleontology* 31:242-272.
723 51. Suter SJ (1994) **Cladistic analysis of cassiduloid echinoids: trying to see the phylogeny for the**
724 **trees** *Biological Journal of the Linnean Society* 53:31-72.
725 52. Kier PM (1966) **Cassiduloids**. In: Moore R, editor. *Treatise on Invertebrate Paleontology, Part U*.
726 Lawrence: University of Kansas Press. pp. 492-523.
727 53. Mortensen T (1948) **A Monograph of the Echinoidea. IV, 1 Holectypoida, Cassiduloida**.
728 Copenhagen: CA Reitzel.
729 54. King B (2020) **Bayesian tip-dated phylogenetics in paleontology: topological effects and**
730 **stratigraphic fit** *Systematic Biology* 70:283-294.
731 55. Turner AH, Pritchard AC, Matzke NJ (2017) **Empirical and Bayesian approaches to fossil-only**
732 **divergence times: a study across three reptile clades** *PLoS One* 12:e0169885.
733 56. Mongiardino Koch N, Garwood RJ, Parry LA (2021) **Fossils improve phylogenetic analyses of**
734 **morphological characters** *Proceedings of the Royal Society B: Biological Sciences* 288:2021004.

735 57. Dos Reis M, Gunnell GF, Barba-Montoya J, Wilkins A, Yang Z, Yoder AD (2018) **Using**
736 **phylogenomic data to explore the effects of relaxed clocks and calibration strategies on divergence**
737 **time estimation: primates as a test case** *Systematic Biology* 67:594-615.
738 58. Warnock RC, Yang Z, Donoghue PC (2011) **Exploring uncertainty in the calibration of the**
739 **molecular clock** *Biology Letters* 8:156-159.
740 59. Dos Reis M, Thawornwattana Y, Angelis K, Telford MJ, Donoghue PC, Yang Z (2015) **Uncertainty**
741 **in the timing of origin of animals and the limits of precision in molecular timescales** *Current Biology*
742 25:2939-2950.
743 60. Sauquet H (2013) **A practical guide to molecular dating**. *Comptes Rendus Palevol* 12:355-367.
744 61. Carruthers T, Scotland RW (2021) **Uncertainty in divergence time estimation** *Systematic Biology*
745 70:855-861.
746 62. Carruthers T, Sanderson MJ, Scotland RW (2020) **The implications of lineage-specific rates for**
747 **divergence time estimation** *Systematic Biology* 69:660-670.
748 63. Lartillot N, Brinkmann H, Philippe H (2007) **Suppression of long-branch attraction artefacts in**
749 **the animal phylogeny using a site-heterogeneous model** *BMC Evolutionary Biology* 7:S4.
750 64. Wang H-C, Minh BQ, Susko E, Roger AJ (2017) **Modeling site heterogeneity with posterior mean**
751 **site frequency profiles accelerates accurate phylogenomic estimation** Systematic Biology 67:216-235.
752 65. Wang H-C, Li K, Susko E, Roger AJ (2008) **A class frequency mixture model that adjusts for site-**
753 **specific amino acid frequencies and improves inference of protein phylogeny** *BMC Evolutionary*
754 *Biology* 8:1-13.
755 66. Li Y, Shen X-X, Evans B, Dunn CW, Rokas A (2021) **Rooting the animal tree of life** *Molecular*
756 *Biology and Evolution* msab170. doi: 10.1093/molbev/msab170.

757 67. Pisani D, Pett W, Dohrmann M, Feuda R, Rota-Stabelli O, Philippe H, et al (2015) **Genomic data**
758 **do not support comb jellies as the sister group to all other animals** *Proceedings of the National*
759 *Academy of Sciences* 112:15402-15407.
760 68. Howard RJ, Puttick MN, Edgecombe GD, Lozano-Fernandez J (2020) **Arachnid monophyly:**
761 **Morphological, palaeontological and molecular support for a single terrestrialization within**
762 **Chelicerata** *Arthropod Structure & Development* 59:100997.
763 69. Ballesteros JA, Sharma PP (2019) **A critical appraisal of the placement of Xiphosura**
764 **(Chelicerata) with account of known sources of phylogenetic error** *Systematic Biology* 68:896-917.
765 70. Whelan NV, Halanych KM (2016) **Who let the CAT out of the bag? Accurately dealing with**
766 **substitutional heterogeneity in phylogenomic analyses** *Systematic Biology* 66:232-255.
767 71. Tao Q, Barba-Montoya J, Huuki LA, Durnan MK, Kumar S (2020) **Relative efficiencies of simple**
768 **and complex substitution models in estimating divergence times in phylogenomics** *Molecular Biology*
769 *and Evolution* 37:1819-1831.
770 72. Smith SA, Brown JW, Walker JF (2018) **So many genes, so little time: A practical approach to**
771 **divergence-time estimation in the genomic era** *PLoS One* 13:e0197433.
772 73. Tao Q, Tamura K, U. Battistuzzi F, Kumar S (2019) **A machine learning method for detecting**
773 **autocorrelation of evolutionary rates in large phylogenies** *Molecular Biology and Evolution* 36:811-824.
774 74. Raup DM, Sepkoski JJ (1982) **Mass extinctions in the marine fossil record** *Science* 215:1501-
775 1503.
776 75. Kier PM (1984) **Echinoids from the Triassic (St. Cassian) of Italy, their lantern supports, and a**
777 **revised phylogeny of Triassic echinoids** *Smithsonian Contributions to Paleobiology* 56:1-41.
778 76. Smith AB (1994) **Triassic echinoids from Peru** *Palaeontographica Abteilung A* 233:177-202.
779 77. Kier PM (1977) **The poor fossil record of the regular echinoid** *Paleobiology* 3:168-174.

780 78. Field D, Berv J, Hsiang A, Lanfear R, Landis M, Dornburg A (2020) **Timing the extant avian**
781 **radiation: the rise of modern birds, and the importance of modeling molecular rate variation** *Bulletin*
782 *of the American Museum of Natural History* 440:150-181.
783 79. Caramés A, Adamonis S, Concheyro A, Remírez M (2019) **New finding of regular echinoid**
784 **elements and microfossils from the Pilmatué Member, Agrio Formation (Early Cretaceous), Neuquén**
785 **Basin, Argentina**. In: Cusminsky G, Bernasconi E, Concheyro G, editors. *Advances in South American*
786 *micropaleontology*. Springer. pp. 1-20.
787 80. R Core Team (2019) **R: A Language and Environment for Statistical Computing**. Vienna, Austria:
788 R Foundation for Statistical Computing.
789 81. Ronquist F, Lartillot N, Phillips MJ (2016) **Closing the gap between rocks and clocks using total-**
790 **evidence dating** *Philosophical Transactions of the Royal Society B* 371:20150136.
791 82. Coiro M, Doyle JA, Hilton J (2019) **How deep is the conflict between molecular and fossil**
792 **evidence on the age of angiosperms?** *New Phytologist* 223:83-99.
793 83. Bolger AM, Lohse M, Usadel B (2014) **Trimmomatic: a flexible trimmer for Illumina sequence**
794 **data** *Bioinformatics* 30:2114-2120.
795 84. Dunn CW, Howison M, Zapata F (2013) **Agalma: an automated phylogenomics workflow** *BMC*
796 *Bioinformatics* 14:330.
797 85. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al (2011) **Full-length**
798 **transcriptome assembly from RNA-Seq data without a reference genome** *Nature Biotechnology*
799 29:644-652.
800 86. Modolo L, Lerat E (2015) **UrQt: an efficient software for the Unsupervised Quality trimming of**
801 **NGS data** *BMC Bioinformatics* 16:1-8.

802 87. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W (2015) **MEGAHIT: an ultra-fast single-node solution**
803 **for large and complex metagenomics assembly via succinct de Bruijn graph** *Bioinformatics* 31:1674-
804 1676.
805 88. Smit A, Hubley R, Green P (2013-2015) **RepeatMasker Open-4.0**.
806 89. Hoff KJ, Stanke M (2019) **Predicting genes in single genomes with AUGUSTUS**. *Current Protocols*
807 *in Bioinformatics*. 65:e57.
808 90. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B (2006) **AUGUSTUS: ab initio**
809 **prediction of alternative transcripts** *Nucleic Acids Research* 34:435-439.
810 91. Simion P, Belkhir K, François C, Veyssier J, Rink JC, Manuel M, et al (2018) **A software tool**
811 **'CroCo'detects pervasive cross-species contamination in next generation sequencing data**. *BMC*
812 *Biology* 16:28.
813 92. Ryan JF (2014) **Alien Index: identify potential non-animal transcripts or horizontally**
814 **transferred genes in animal transcriptomes**. doi: 10.5281/zenodo.21029.
815 93. Guang A, Howison M, Zapata F, Lawrence C, Dunn CW (2021) **Revising transcriptome**
816 **assemblies with phylogenetic information** *PLos One* 16:e0244202.
817 94. Ranwez V, Harispe S, Delsuc F, Douzery EJ (2011) **MACSE: Multiple Alignment of Coding**
818 **SEquences accounting for frameshifts and stop codons** *PLoS One* 6:e22594.
819 95. Talavera G, Castresana J (2007) **Improvement of phylogenies after removing divergent and**
820 **ambiguously aligned blocks from protein sequence alignments** *Systematic Biology* 56:564-577.
821 96. Morel B, Kozlov AM, Stamatakis A (2018) **ParGenes: a tool for massively parallel model**
822 **selection and phylogenetic tree inference on thousands of genes** *Bioinformatics* 35:1771-1773.
823 97. Mai U, Mirarab S (2018) **TreeShrink: fast and accurate detection of outlier long branches in**
824 **collections of phylogenetic trees** *BMC Genomics* 19:272.

825 98. Zhang C, Rabiee M, Sayyari E, Mirarab S (2018) **ASTRAL-III: polynomial time species tree**
826 **reconstruction from partially resolved gene trees** *BMC Bioinformatics* 19:153.
827 99. Sayyari E, Mirarab S (2016) **Fast coalescent-based computation of local branch support from**
828 **quartet frequencies** *Molecular Biology and Evolution* 33:1654-1668.
829 100. Aberer AJ, Kobert K, Stamatakis A (2014) **ExaBayes: massively parallel Bayesian tree inference**
830 **for the whole-genome era** *Molecular Biology and Evolution* 31:2553-2556.
831 101. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ (2014) **IQ-TREE: a fast and effective**
832 **stochastic algorithm for estimating maximum-likelihood phylogenies** *Molecular Biology and Evolution*
833 32:268-274.
834 102. Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A, Jermiin LS (2017) **ModelFinder: fast**
835 **model selection for accurate phylogenetic estimates** *Nature Methods* 14:587.
836 103. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS (2017) **UFBoot2: improving the**
837 **ultrafast bootstrap approximation** *Molecular Biology and Evolution* 35:518-522.
838 104. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A (2019) **RAxML-NG: a fast, scalable and**
839 **user-friendly tool for maximum likelihood phylogenetic inference** *Bioinformatics* 35:4453-4455.
840 105. Wang H-C, Minh BQ, Susko E, Roger AJ (2018) **Modeling site heterogeneity with posterior mean**
841 **site frequency profiles accelerates accurate phylogenomic estimation** *Systematic Biology* 67:216-235.
842 106. Lartillot N, Rodrigue N, Stubbs D, Richer J (2013) **PhyloBayes MPI. Phylogenetic reconstruction**
843 **with infinite mixtures of profiles in a parallel environment** *Systematic Biology* 62:611-615.
844 107. Strimmer K, Von Haeseler A (1997) **Likelihood-mapping: a simple method to visualize**
845 **phylogenetic content of a sequence alignment** *Proceedings of the National Academy of Sciences*
846 94:6815-6819.
847 108. Shen X-X, Hittinger CT, Rokas A (2017) **Contentious relationships in phylogenomic studies can**
848 **be driven by a handful of genes** *Nature Ecology &* Evolution 1:0126.

849 109. Yang Z, Rannala B (2006) **Bayesian estimation of species divergence times under a molecular**
850 **clock using multiple fossil calibrations with soft bounds** *Molecular Biology and Evolution* 23:212-226.
851 110. Telford MJ, Lowe CJ, Cameron CB, Ortega-Martinez O, Aronowicz J, Oliveri P, et al (2014)
852 **Phylogenomic analysis of echinoderm class relationships supports Asterozoa** *Proceedings of the Royal*
853 *Society of London B: Biological Sciences* 281:20140479.
854 111. Schlager S (2017) **Morpho and Rvcg–Shape Analysis in R**. In: Zheng G, Li S, Székely GJ, editors.
855 *Statistical shape and deformation analysis*. Elsevier. pp. 217-256.
856 112. Paradis E, Schliep K (2018) **ape 5.0: an environment for modern phylogenetics and**
857 **evolutionary analyses in R** *Bioinformatics* 35:526-528.
858

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 [113]. 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

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

883 The assembled transcriptomes were sanitized from cross-contaminant reads product of
884 multiplexed sequencing using CroCo v. 1.1 [91]. These eight datasets, as well as two annelid
885 transcriptomes not employed in this study but sequenced in the same lane, were incorporated.
886 Transcripts considered over or under-expressed across samples (as defined by default parameters),
887 were kept. This resulted in an average removal of 1.26% of assembled transcripts (range: 0.4% for
888 *Aspidodiadema* to 3.06% for *Histocidaris*).

889

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

899

900 *Bathysalenia phoinissa*, *Micropyga tuberculata*: The material was collected by the deep-sea
901 cruise BIOMAGLO [114] conducted in 2017 jointly by the French National Museum of Natural History
902 (MNHN) as part of the Tropical Deep-Sea Benthos program, the French Research Institute for
903 Exploitation of the Sea (IFREMER), the "Terres Australes et Antarctiques Françaises" (TAAF), the

904 Departmental Council of Mayotte and the French Development Agency (AFD), with the financial support
905 of the European Union (Xe FED). Specimens were sorted into taxonomic bins at class level and
906 collectively fixed in 70% ethanol at room temperature. Following taxonomic identification, tissue
907 samples were obtained using sterilized forceps and disposable scalpel blades. Tissue subsamples were
908 collected in high-purity 96% ethanol and stored at -40°C until DNA extraction. Total genomic DNA was
909 extracted using the DNeasy Blood and Tissue Kit (QIAGEN) following the manufacturer's instructions,
910 complemented by the addition of 4 µl RNAse A (QIAGEN, 100 mg/ml) for RNA digestion. Elution buffer
911 volume was reduced to 60 µl and the elution step repeated twice per sample in order to maximize DNA
912 concentration and yield. After collection of the first elution, spin-column membranes were rinsed again
913 twice with fresh elution buffer (40 µl) to further increase yield. Elutions were collected separately, with
914 the first one used for sequencing and the second for quality control. For the latter, a portion of the
915 mitochondrial cytochrome c oxidase subunit I (COI) was amplified using PCR. Amplifications were
916 conducted with TopTaq DNA Polymerase (QIAGEN) using 1 µl of extracted genomic DNA (approx. 10–15
917 ng). Details on primers and protocols can be found in [115]. PCR products were visualized on a 1%
918 agarose gel, purified using ExoSAP-IT (Affymetrix), and sequenced at Microsynth GmbH (Vienna, Austria)
919 with the same primers. Sequences are deposited in GenBank under the accession numbers XXX and XXX.

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

928 (https://sourceforge.net/projects/bbmap/) with settings: `minlen=50 ktrim=r k=23 mink=11`
929 `tpe tbo`); followed by quality trimming of reads using UrQt v. 1.0.18 [86], discarding regions with
930 phred quality score < 28 and sequences of size < 50. Assembly of the data was carried out with MEGAHIT
931 v. 1.1.2 [87], in an iterative approach followed by assessment of assembly quality using BUSCO v. 3.0.2.
932 scores [116] using the metazoan dataset ODB v. 9 [117]. Final BUSCO scores of the draft assemblies
933 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%,
934 F:59.0%, M:23.8%, n:978) for *Micropyga*. As recommended by Hoff and Stanke [89], draft genomes were
935 masked with RepeatMasker v. 4.1.0 [88] prior to gene prediction, using settings: `-norna -xsmall`;
936 resulting in 3.82 % and 2.9% masked bases for *Bathysalenia* and *Micropyga*, respectively. Gene
937 prediction was carried out using AUGUSTUS v. 3.2.3 [90] using settings: `--strand=both --`
938 `singlestrand=false --genemodel=partial --codingseq=on --sample=0 --`
939 `alternatives-from-sampling=false --exonnames=on --softmasking=on`. This step
940 relied on a training dataset composed of universal single-copy orthologs (USCOs) derived from the most
941 recent *Strongylocentrotus purpuratus* genome (Spur v.5).

942

943 *Echinothrix calamaris*, *Tripneustes gratilla*. Tissue samples were taken from live sea urchins
944 housed in laboratory aquaria and total RNA was immediately extracted from 150 mg of tissue using a
945 PureLink™ RNA extraction kit (Invitrogen™). The quality of total RNA was assessed on a BioAnalyzer
946 2100 (Agilent Technologies, Santa Clara, CA, USA) to ensure a RIN > 9 for all samples. Mature mRNA was
947 extracted from 1ug of total RNA and cDNA libraries were constructed using a TruSeq kit (Illumina).
948 Quality control of libraries was assessed on a BioAnalyzer and quantification measured using qPCR.
949 *NEBNext* Multiplex adaptors were added via ligation, and the cDNA libraries were sequenced at Genome
950 Quebec, McGill University. Three libraries were multiplex sequenced on one lane of Illumina at a

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

953

954 *Tetrapygus niger*. Adult specimens were collected and transported alive to the University of
955 Concepcion. Tissue samples (female gonad and tube feet) from one individual were finely chopped with
956 a scalpel, and total RNA was extracted immediately using TriReagent (Sigma), following the
957 manufacturer's instructions. Total RNA concentration was estimated using QuantiFluor® RNA System
958 (111.84-339.04 ng/µL) and quality was assessed using Agilent RNA 6000 Nano kit with an Agilent 2100
959 TapeStation (RIN: 5.4-8.4). cDNA libraries were prepared at Genoma Mayor SpA (Chile) using TruSeq
960 Stranded mRNA Kit (Illumina) and sequenced on a multiplexed run using Illumina HiSeq 4000 platform
961 with 150 paired-end reads, resulting in 54.1 M reads for gonad female and 52.7 M reads for tube feet,
962 respectively. These two transcriptomic datasets were combined before performing all steps of
963 bioinformatic processing.

964

965 *Eucidaris metularia*. A single specimen was preserved in RNA*later* after being starved overnight
966 and kept for 1 day at 4°C before long-term storage at -80°C. Total RNA was extracted using Direct-zol
967 RNA Miniprep Kit (with in-column DNase treatment; Zymo Research) from Trizol. mRNA was isolated
968 with Dynabeads mRNA Direct Micro Kit (Invitrogen). RNA concentration was estimated using Qubit RNA
969 broad range assay kit, and quality was assessed using RNA ScreenTape with an Agilent 4200 TapeStation.
970 Values were used to customize downstream protocols following manufacturers' instructions. Library
971 preparation was performed with KAPA-Stranded RNA-Seq kits, targeting an insert size in the range of
972 200–300 base pairs. The quality and concentration of libraries were assessed using DNA ScreenTape.
973 The library was sequenced in a multiplexed pair-end runs using Illumina HiSeq 4000 with 7 other

974 libraries in the same lane (all previously published in [4]). In order to minimize read crossover, 10 bp
975 sequence tags designed to be robust to indel and substitution errors were employed [113].

976 The assembled transcriptome was filtered from cross-contaminant reads product of multiplexed
977 sequencing using CroCo v. 1.1 [91]. Seven other transcriptomes sequenced together were also
978 employed, and results of this step are reported in Fig. S1 of Mongiardino Koch & Thompson [3].

979

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:** [118]. **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**: [118]. **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 [43]. The first records of stereom point to a sudden and global appearance within the middle Atdabanian [42]. 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:** [119, 120]. **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**: [119, 120]. **Taxon**:

1002 *Maydenia roadsidensis.* **Notes**: The divergence of asterozoan classes is calibrated based upon the
1003 stratigraphic occurrence of *Maydenia roadsidensis*, the oldest asterozoan, which is a stem group
1004 ophiuroid [121] and occurs in the *Psigraptus jacksoni* Zone of the Floian.

1005 Asteroidea – **Younger bound:** 239.10 Ma, Top Fassanian, Ladinian, Mo3, *nodosus* biozone, Triassic.
1006 **Older bound**: 252.16 Ma, P-T Boundary. **Reference**: [122]. **Taxon**: *Trichasteropsis weissmanni*. **Notes**:
1007 Fossil evidence suggests the divergence of the asterozoan crown group took place sometime in the Early
1008 or Middle Triassic [122]. Based upon its phylogenetic position and stratigraphic occurrence as the oldest
1009 crown group asteroid, we use the forcipulatacean *T. weissmanni* from the Middle Triassic Muschelkalk
1010 as the soft bound on this divergence.

1011 Crinoidea – **Younger bound**: 247.06 Ma, Top Spathian (Paris Biota, Lower Shale unit, Thaynes Group,
1012 Spathian, Triassic). **Older bound**: Unconstrained. **Reference**: [123]. **Taxon**: *Holocrinus*. **Notes**: Crown
1013 group Crinoidea is difficult to define, though a rapid post-Palaeozoic morphological diversification is
1014 supported by fossil evidence. We use *Holocrinus* from the Early Triassic of the Thaynes group to calibrate
1015 the younger bound on the divergence of the crown group (i.e., the split between Isocrinida and all other
1016 extant crinoids).

1017 Ophiuroidea (Euryophiurida-Ophintegrida divergence) – **Younger bound**: 247.06 Ma, Top Spathian (Paris
1018 Biota, Lower Shale unit, Thaynes Group, Spathian, Triassic). **Older bound**: Unconstrained. **Reference**:
1019 [124]. **Taxon**: *Shoshonura brayardi*. **Notes**: *S. brayardi* from the Spathian Thaynes Group of the Western
1020 USA is a member of the crown group ophiuroid clade Ophiodermatina. It is thus the currently oldest
1021 described member of the ophiuroid crown group, setting its minimum age of divergence.

1022 Pneumonophora (Holothuriida-Neoholothuriida divergence) – **Younger bound**: 259.8 Ma, S*pinosus* zone
1023 of Early Ladinian (used base Ladinian). **Older bound**: 241.5 Ma, Base Wuchiapingian. **Reference**: [46,

119]. **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**: [125]. **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. [126]. **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.

1046 Stomopneustoida-Camarodonta divergence – **Younger bound**: 201.3 Ma, Top Pliensbachian. **Older**
1047 **bound**: 182.7 Ma, Base Jurassic. **Reference**: See discussion in supplement of Thompson et al. [13].
1048 **Taxon**: Stomechinids from Morocco like *Diplechinus hebbriensis* [127]. **Notes**: The oldest stomechinids,
1049 which are stem group stomopneustoids, are from the Early Jurassic of Morocco.

1050 Neognathostomata-Atelostomata divergence – **Younger bound**: 234.5 Ma, Base of Toarcian. **Older**
1051 **bound**: 182.7 Ma, St. Cassian Beds, bottom of Julian 2 ammonoid Zone. **Reference**: [14, 126]. **Taxon**:
1052 Younger bound is set by *Galeropygus sublaevis* (older bound is relatively uninformative). **Notes**: The
1053 base of the Toarcian (which contains the oldest neognathostomate *G. sublaevis*) is used as the lower
1054 bound on the divergence between the neognathostomata and all other crown irregular echinoids. Note
1055 that depending on the position of the unsampled echinoneoids, this split might also correspond to the
1056 origin of crown group Irregularia (see recent topologies recovered by Lin et al. [36] and Mongiardino
1057 Koch & Thompson [3]).

1058 Atelostomata (Holasteroida-Spatangoida divergence) – **Younger bound**: 137.68 Ma, Bottom of
1059 *Campylotoxus* zone in Valanginian. **Older bound**: 201.3 Ma, Base of Jurassic. **Reference**: [128]. **Taxon**:
1060 Younger bound is based on *Toxaster* (older bound is loosely uninformative). **Notes**: The toxasterids are
1061 stem group spatangoids that calibrate the divergence between spatangoids and holasteroids. The oldest
1062 *Toxaster* are from the *Campylotoxus* zone in the Valanginian, and set the younger bound on the
1063 divergence.

1064 Echinolampadacea (Cassiduloida-Scutelloida divergence) – **Younger bound**: 113 Ma, Base Albian. **Older**
1065 **bound**: 145 Ma, Base Cretaceous. **Reference**: [129]. **Taxon**: Younger bound is based on *Eurypetalum*
1066 *rancheriana* (older bound is loosely uninformative). **Notes**: The oldest member of Echinolampadacea is
1067 the faujasiid *Eurypetalum rancheriana* (originally placed in the genus *Faujasia* and later transferred, see
1068 [28, 29]) from the Late Albian of Colombia.

Scutelloida (Scutelliformes-Laganiformes divergence) – **Younger bound**: 56.0 Ma, Bottom of Eocene. **Older bound**: Unconstrained. **Reference**: [14, 130]. **Taxon**: Younger bound is based on *Eoscutum doncieuxi*. **Notes**: We use the base of the Eocene as the younger bound on the divergence between scutelliforms and laganiforms. This is based upon the occurrence of *E. doncieuxi*, the oldest scutelliform, which is known from the Lower Eocene. The older bound is left unconstrained in order to account for the current uncertainty in the origin of the clade (i.e., allow for Mesozoic ages).

Leodia-(*Encope*+*Mellita*) divergence – **Younger bound**: 23.03 Ma, Base Miocene. **Older bound**: 33.9 Ma, Base of Oligocene. **Reference**: [131, 132]. **Taxon**: Younger bound is based on *Encope michoacanensis* (older bound is loosely uninformative). **Notes**: This node is constrained using the oldest representative of these three genera of mellitids.

Clypeasteroida – **Younger bound**: 47.8 Ma, Base Lutetian. **Older bound**: Unconstrained. **Reference**: [133, 134]. **Taxon**: Younger bound is based on *Clypeaster calzadai* and *Clypeaster moianensis*. **Notes**: The oldest known fossil species of Clypeasteroida (the clypeasterids *C. calzadai* and *C. moianensis*) are from the middle Eocene, upper Lutetian (Biarritzian) of Cataluña, Spain. The older bound is left unconstrained in order to account for the current uncertainty in the origin of the clade (i.e., allow for Mesozoic ages).

Strongylocentrotidae-Toxopneustidae divergence – **Younger bound**: 44.4 Ma, Eocene Planktonic Zones E9-E12. **Older bound**: 56.0 Ma, Base Eocene. **Reference**: [135, 136]. **Taxon**: Younger bound is based on *Lytechinus axiologus* (older bound is loosely uninformative). **Notes**: The oldest member of the clade comprising toxopneustids and strongylocentrotids is the toxopneustid *Lytechinus axiologus.* This taxon is known from the Eocene Yellow Limestone Group of Jamaica. The Yellow Limestone group represents the Eocene Planktonic Zones E9-E12.

Supplementary Figures and Tables

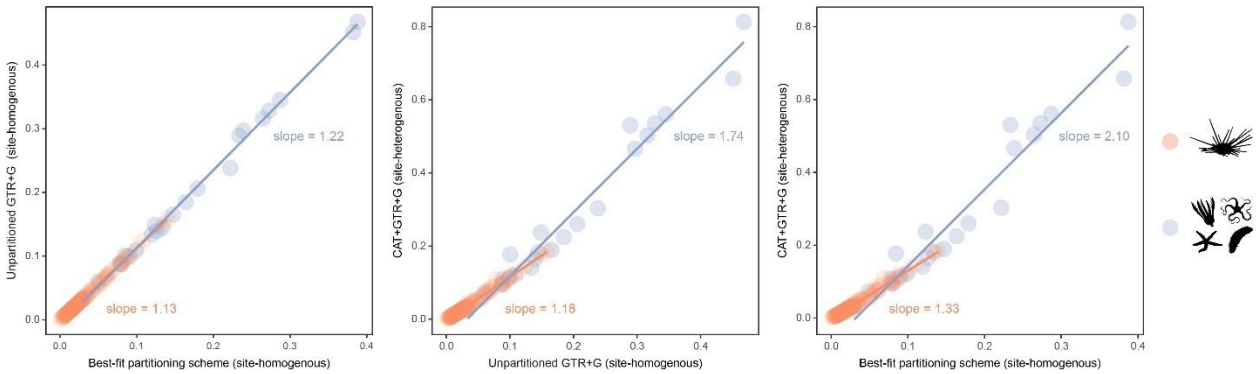

Figure S1: Estimated branch lengths across different models of molecular evolution. Different site-homogeneous models (left) infer similar levels of divergence, and the choice between them induces little distortion in the general tree structure. Site-heterogeneous models on the other hand, not only infer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

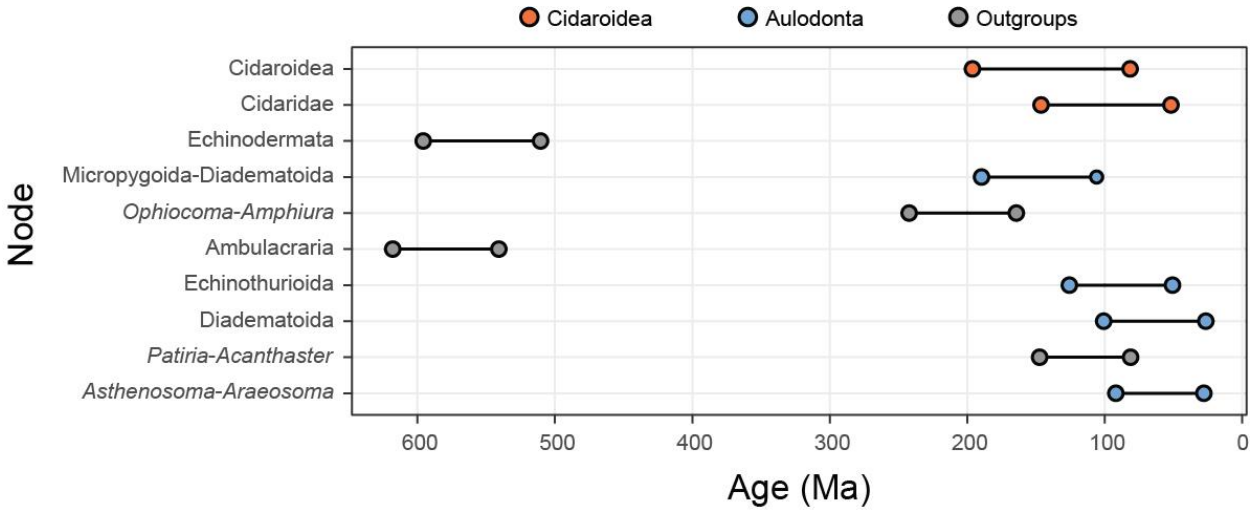

Figure S2: The ten most sensitive node dates are found within Cidaroidea, Aulodonta, and among outgroup nodes. For each, the range shown spans the interval between the minimum and maximum ages found among the consensus topologies of all time-calibrated runs.

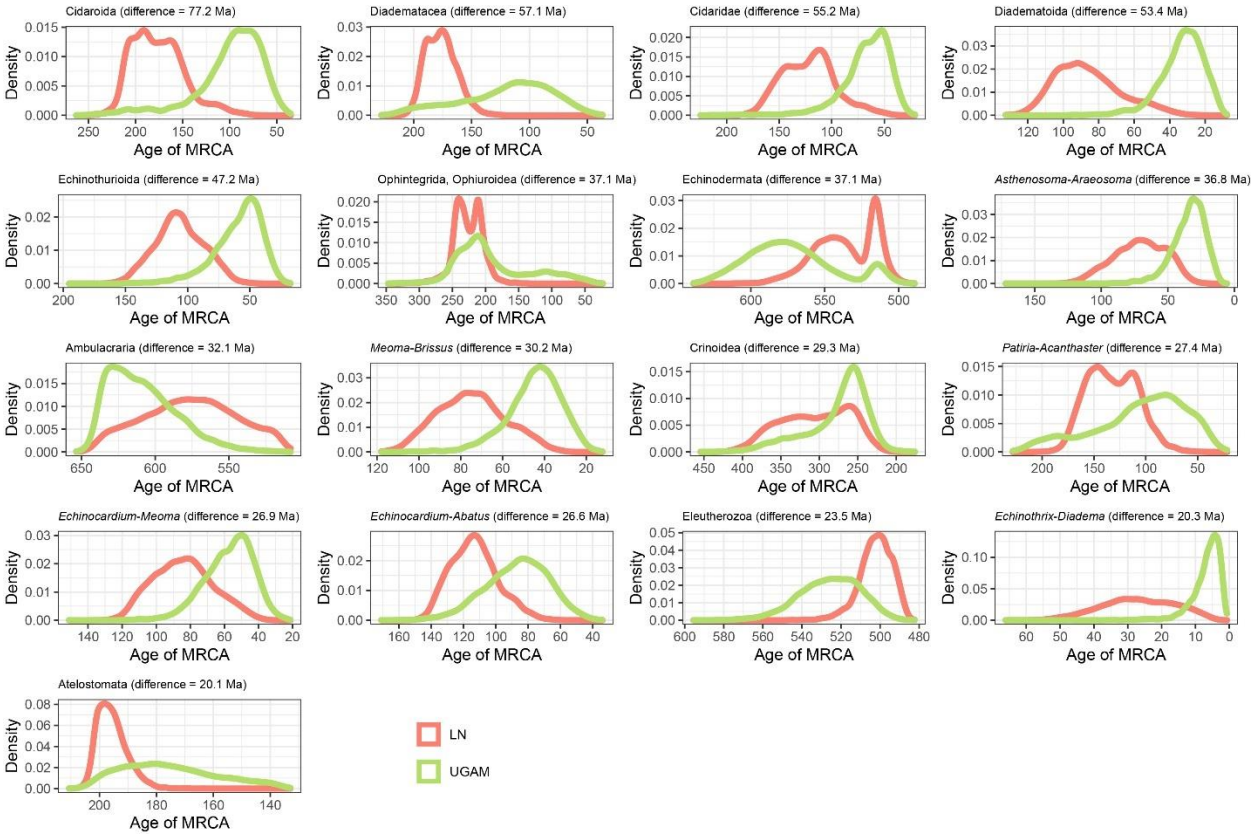

Figure S3: Distribution of posterior probabilities for node ages that show an average difference larger than 20 Ma depending on the choice of clock prior.

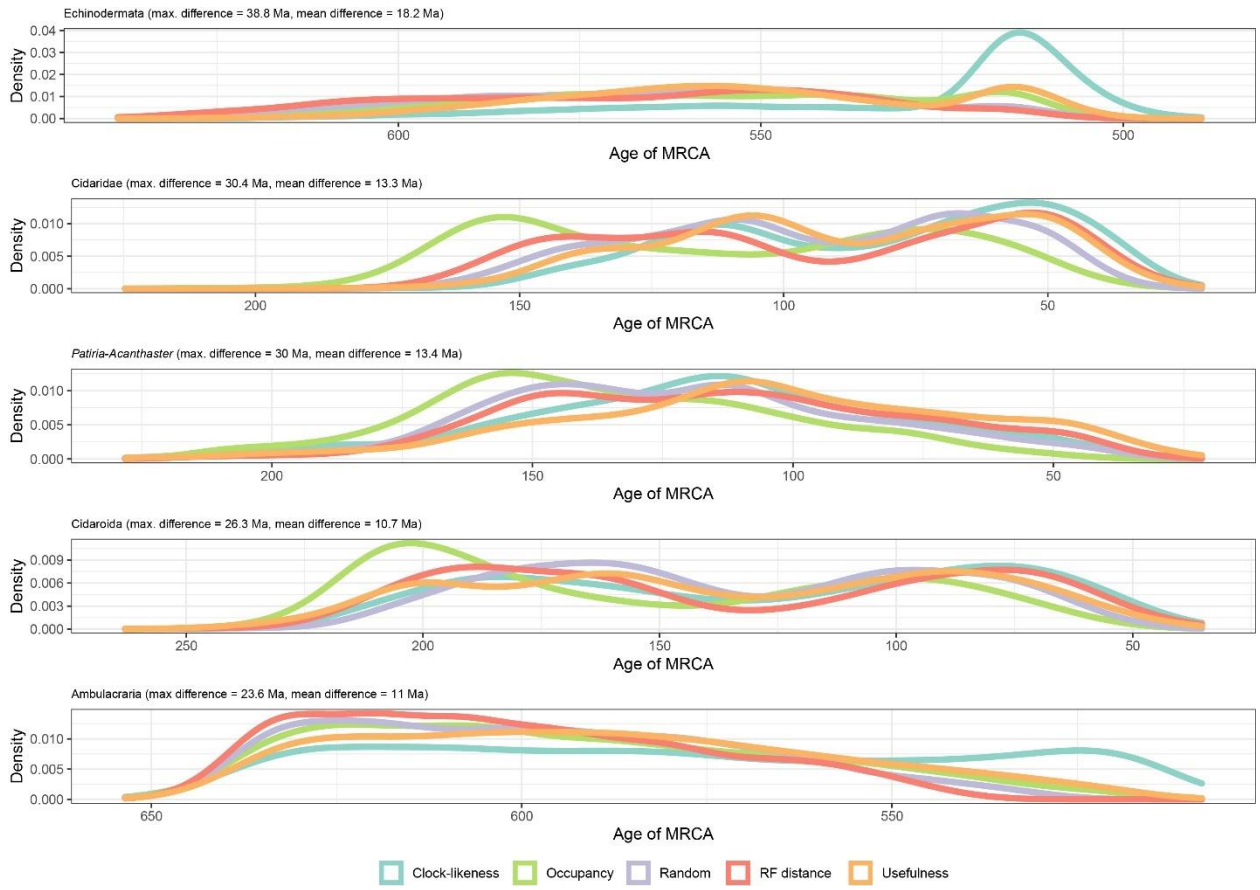

Figure S4: Distribution of posterior probabilities for node ages that show a maximum difference larger than 20 Ma depending on the choice of clock prior. The largest differences can be seen in the relatively younger ages of Ambulacraria and Echinodermata when using clock-like genes, and in the relatively 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. 3C as short distances between their centroids in chronospace.

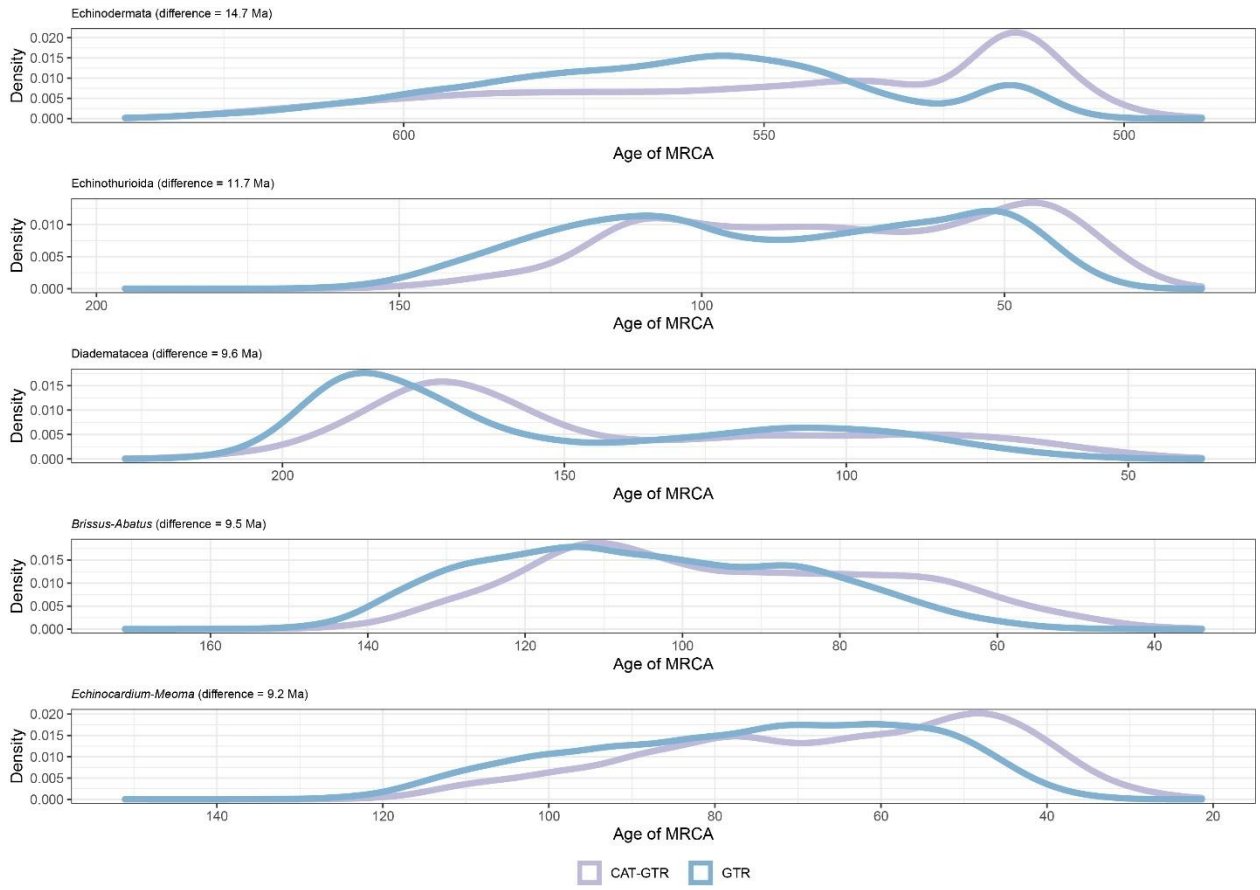

Figure S5: Distribution of posterior probabilities for node ages that are the most affected by the choice of model of molecular evolution. No node showed average differences larger than 20 Ma, so those with the largest effect are shown.

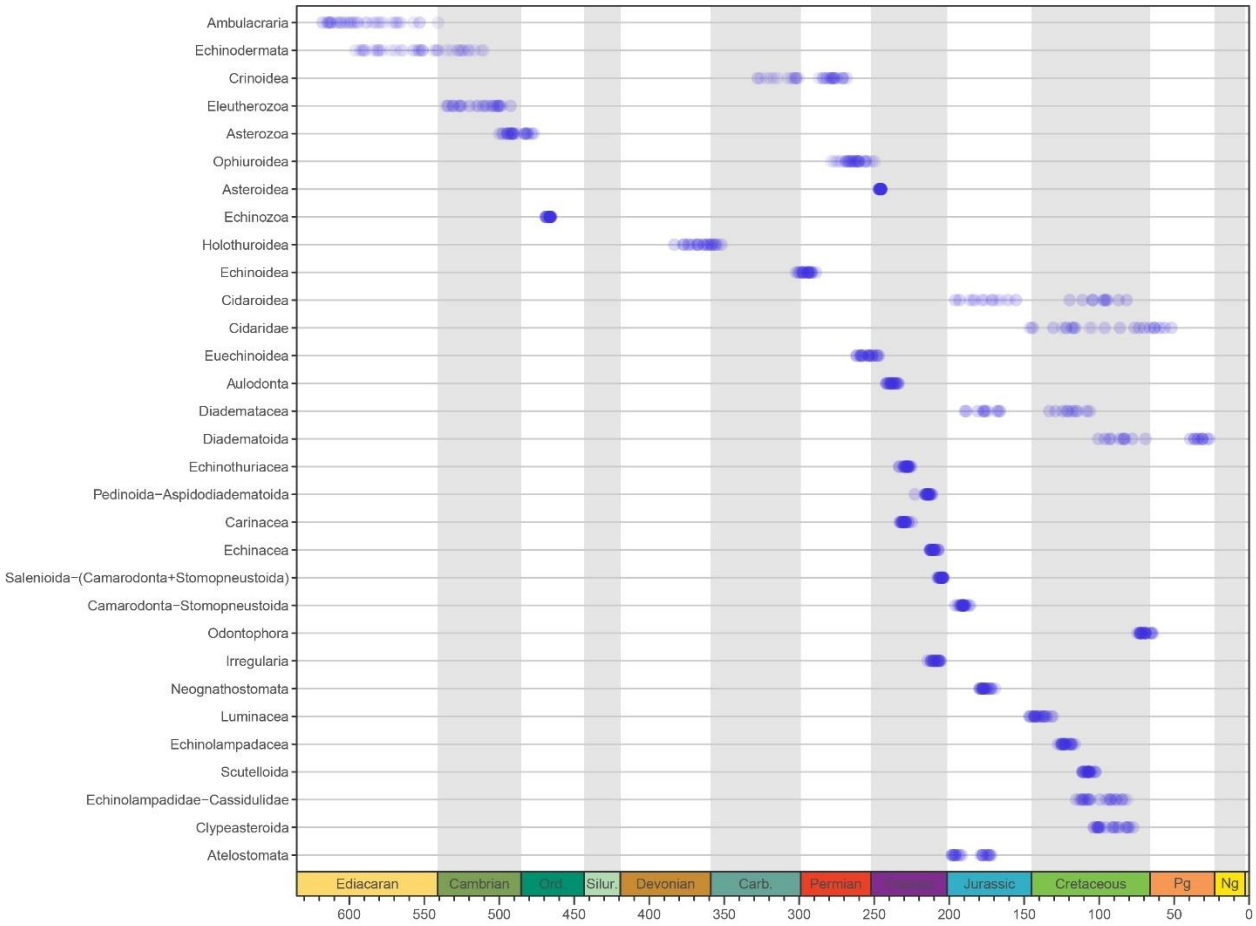

Figure S6: Median ages for selected clades across the consensus trees of the 40 time-calibrated experiments performed (20 analyses x 2 runs each).

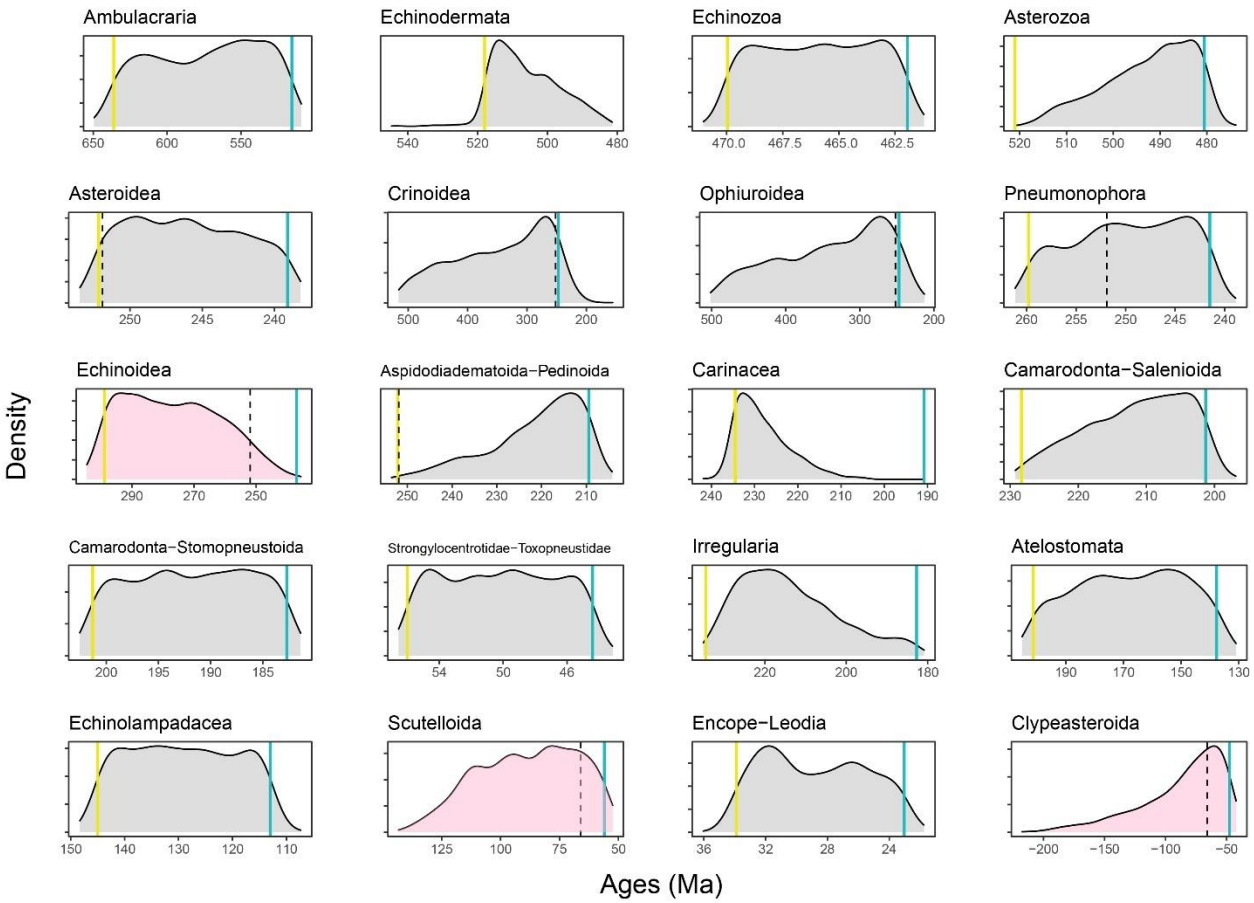

Figure S7: Prior distributions of all constrained nodes. Five hundred replicates were sampled from the joint prior, showing appropriately broad distributions of node ages. Blue lines show minima and yellow ones maxima (when enforced); dotted lines show the age of the Permian-Triassic (251.9 Ma) and Cretaceous-Paleogene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

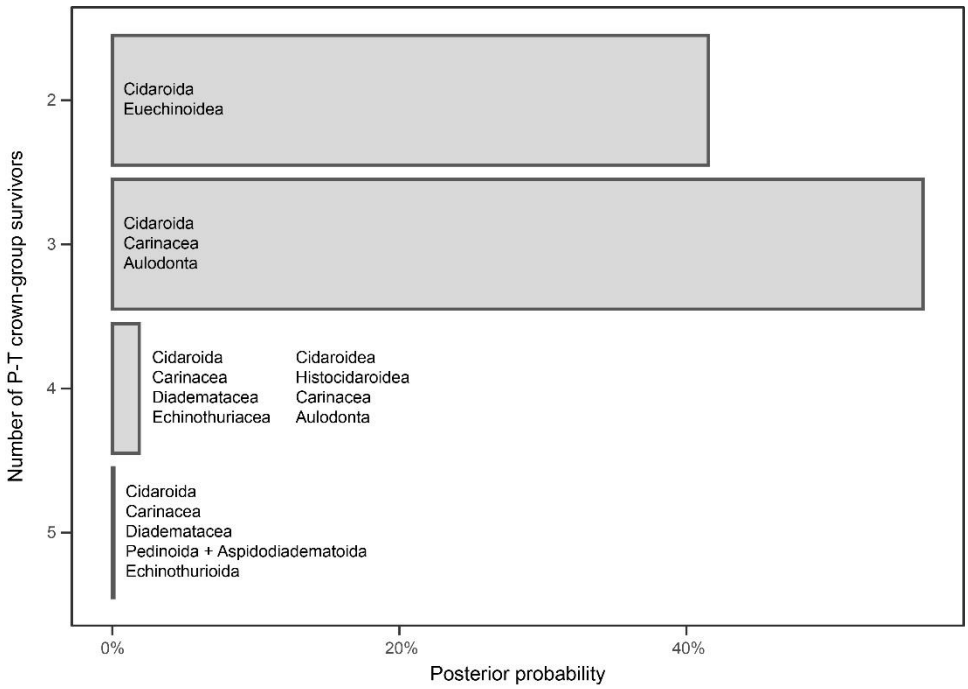

Figure S8: Number of lineages inferred to have crossed the Permian-Triassic (P-T) boundary. The probabilities of each scenario are estimated from the inferred divergence times of the 10,000 chronograms sampled across all of the analyses performed (250 for the two runs of each combination of sampled loci,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

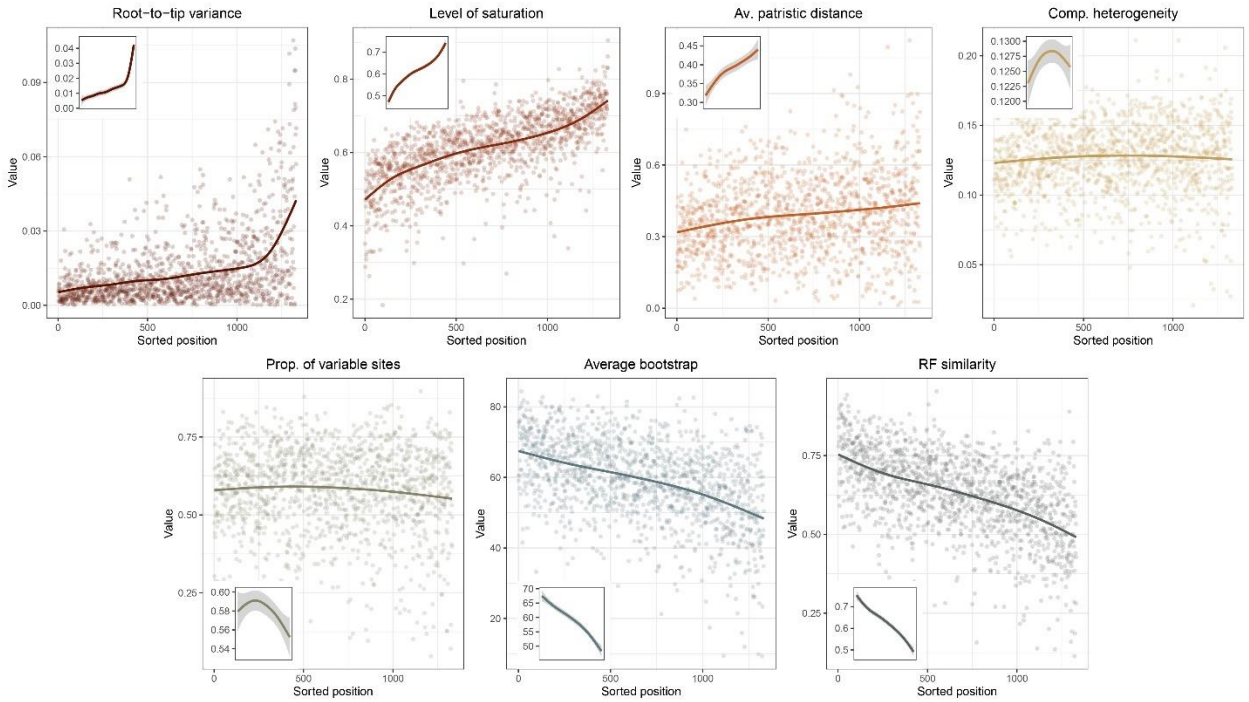

Figure S9: Ordering of loci enforced using *genesortR* [37] and its relationship to the seven gene properties employed. High ranking loci (i.e., the most phylogenetically useful) show low root-to-tip variances (or high clock-likeness), low saturation and compositional heterogeneity, as well as high average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

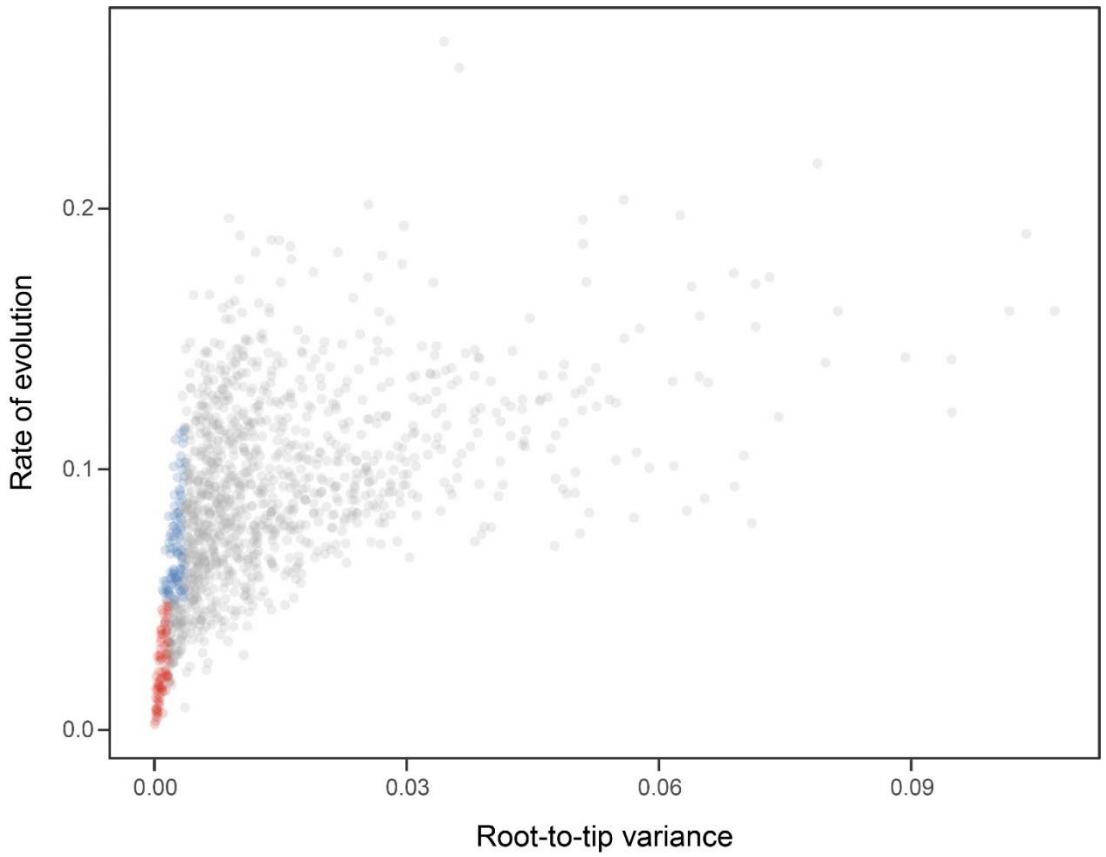

Figure S10: Relationship between the root-to-tip variance (a proxy for the clock-likeness of loci) and the rate of evolution (estimated as the total tree length divided by the number of terminals). The most clock-like loci (shown in red), which are often favored for the inference of divergence times (e.g., [62, 72]), 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).

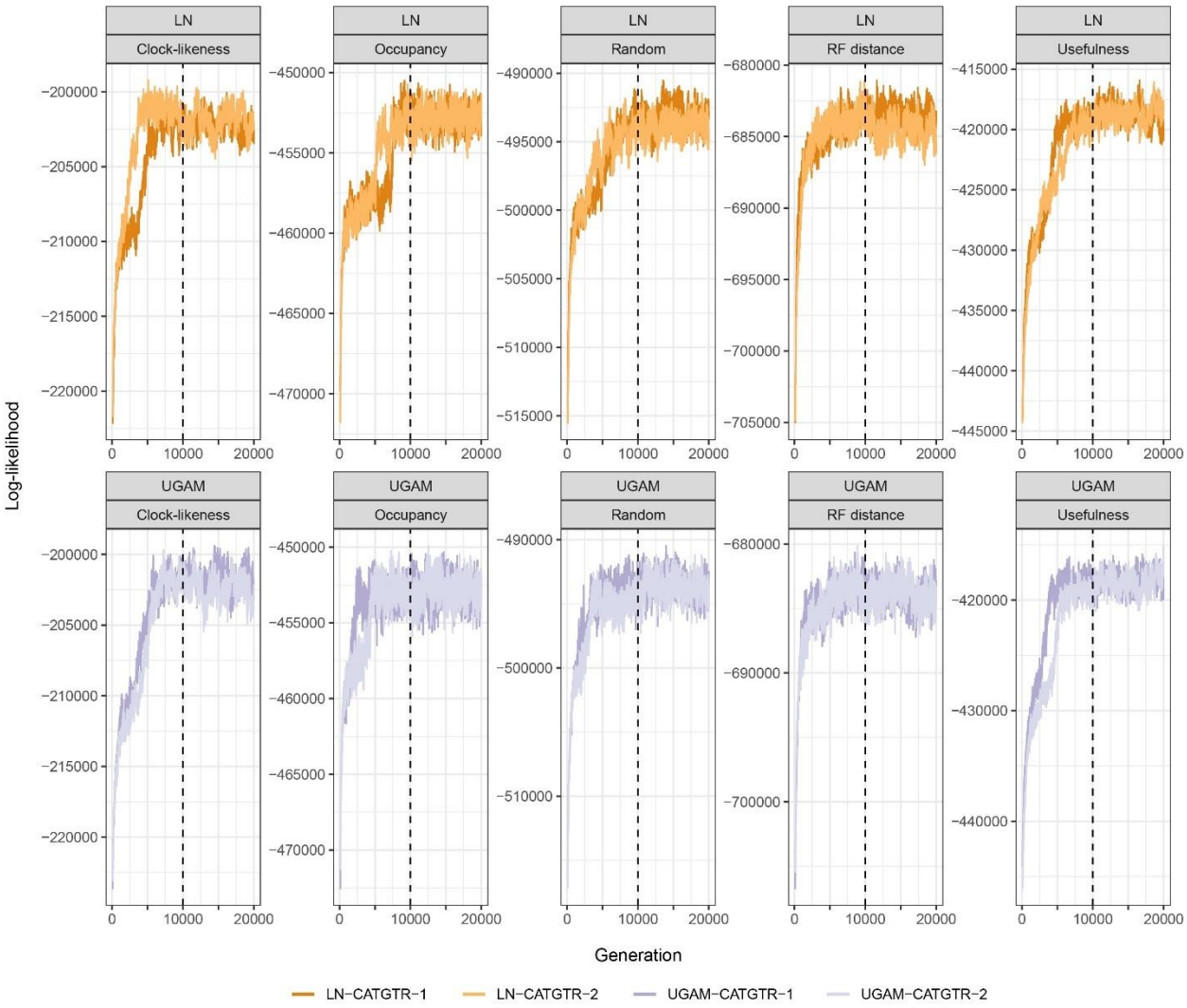

Figure S11: Trace plots of the log-likelihood of different time-calibration runs. All runs show evidence of reaching convergence and stationarity before our imposed burn-in fraction of 10,000 generations. For simplicity, only runs under the CAT+GTR+G model are plotted. Those run under GTR+G converged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

1152 **Table S1:** Transcriptomic/genomic datasets added in this study relative to the taxon sampling of
1153 Mongiardino Koch et al. [4] and Mongiardino Koch & Thompson [3]. Taxa with citations were taken from
1154 the literature, and details can be found in the corresponding studies and associated NCBI BioProjects.
1155 Taxa are shown in alphabetical order; those identified with "OG" are outgroup taxa (i.e., non-echinoids).
1156 Voucher specimens are deposited at the Benthic Invertebrate Collection, Scripps Institution of
1157 Oceanography (SIO-BIC), and the Echinoderm Collection, Muséum National d'Histoire Naturelle (MNHN-
1158 IE). If multiple accession numbers are shown for a given taxon, these datasets were merged before
1159 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)	[137]	-	-	-	-	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)	[138]	-	-	-	-	SRX445860
Astrophyton muricatum (OG)	[139, 140]	-	-	-	-	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)	[140, 141]	-	-	-	-	SRX1397823
Leodia sexiesperforata	This study	Eggs	Frances Armstrong	Bocas del Toro, Panama	-	
Metacrinus rotundus (OG)	[142]	-	-	-	-	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)	[138]	-	-	-	-	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	[143]	-	-	-	-	DRX187887 DRX187888
Sinaechinocyamus mai	This study	Gregory's diverticulum	Jih-Pai Lin	Miaoli County, Taiwan	SIO-BIC E7919	
Sterechinus neumayeri	[144]	-	-	-	-	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

1160

1161 **Table S2:** Details of molecular datasets and supermatrix. Statistics for raw reads and assemblies are
1162 shown for datasets incorporated in this study relative to the sampling of [3, 4] (where similar statistics
1163 can be found for the other datasets). Taxa are shown in alphabetical order; those identified with "OG"
1164 are outgroup taxa (i.e., non-echinoids). Novel datasets correspond to Illumina pair-end transcriptomes,
1165 except for two draft genomes (*Bathysalenia phoinissa* and *Micropyga tuberculata*). Mean insert size is
1166 expressed in number of base pairs. For transcriptomes, read pairs (initial) shows numbers input into
1167 Agalma [84], (i.e., after processing with Trimmomatic [83] or UrQt [86]), while read pairs (retained)
1168 shows those that passed the Agalma curation checks (including ribosomal, adapter, quality, and
1169 compositional filters), and represent the final number of read pairs used for assembly. For genomes, see
1170 information in the bioinformatic details above. Assemblies were further sanitized with either
1171 alien_index [92] alone or in combination with CroCo [91] (see bioinformatic details), and the number of
1172 assembled transcripts/contigs shows the size of datasets after these curation steps. The number of loci
1173 shows the occupancy of terminals in the supermatrix output by Agalma (1,346 loci at 70% occupancy),
1174 after which loci were further removed by TreeShrink [97], 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

1175

1176 ## Supplementary References (references 1-112 in main text)

1177 113. Faircloth BC, Glenn TC (2012) **Not all sequence tags are created equal: designing and validating**
1178 **sequence identification tags robust to indels** *PLoS One* 7:e42543.

1179 114. BIOMAGLO cruise, RV Antea (2017). doi: 10.17600/17004000.

1180 115. Bronstein O, Kroh A, Miskelly AD, Smith SD, Dworjanyn SA, Mos B, et al (2019) **Implications of**
1181 **range overlap in the commercially important pan-tropical sea urchin genus *Tripneustes* (Echinoidea:**
1182 **Toxopneustidae)** *Marine Biology* 166:34.

1183 116. Seppey M, Manni M, Zdobnov EM (2019) **BUSCO: assessing genome assembly and annotation**
1184 **completeness** *Methods in Molecular Biology* 1962:227-245.

1185 117. Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simao FA, Pozdnyakov IA, et al (2015)
1186 **OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software** *Nucleic*
1187 *acids research* 43:250-256.

1188 118. Benton MJ, Donoghue PC, Asher RJ, Friedman M, Near TJ, Vinther J (2015) **Constraints on the**
1189 **timescale of animal evolutionary history** *Palaeontologica Electronica* 18:1-106.

1190 119. Erkenbrack EM, Thompson JR (2019) **Cell type phylogenetics informs the evolutionary origin of**
1191 **echinoderm larval skeletogenic cell identity** *Communications Biology* 2:160.

1192 120. Cooper R, Sadler P, Hammer O, Gradstein F (2012) **The Ordovician Period**. In: Gradstein FM, Ogg
1193 JG, Schmitz M, Ogg G, editors. *The geologic time scale 2012*. Elsevier. pp. 489-523.

1194 121. Hunter AW, Ortega-Hernández J (2021) **A new somasteroid from the Fezouata Lagerstätte in**
1195 **Morocco and the Early Ordovician origin of Asterozoa** *Biology Letters* 17:20200809.

1196 122. Villier L, Brayard A, Bylund KG, Jenks JF, Escarguel G, Olivier N, et al (2018) ***Superstesaster***
1197 ***promissor* gen. et sp. nov., a new starfish (Echinodermata, Asteroidea) from the Early Triassic of Utah,**
1198 **USA, filling a major gap in the phylogeny of asteroids** *Journal of Systematic Palaeontology* 16:395-415.

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

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

125. Kroh A (2011) **Echinoids from the Triassic of St. Cassian-A review** *Geo.Alp* 8:136-140.

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

127. Smith AB (2011) ***Gymnodiadema* and the Jurassic roots of the Arbacioida (stirodont echinoids)** *Swiss Journal of Palaeontology* 130(1):155-171.

128. François É, Marchand D, David B (2003) **Fluctuations morphologiques et hétérochronies chez Toxaster (échinides, Crétacé inférieur)** *Comptes Rendus Palevol* 2:597-605.

129. Cooke CW (1955) **Some Cretaceous echinoids from the Americas** *US Geological Survey Professional Papers* 264:87-112.

130. Roman J (1989) **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. pp. 18-21.

131. Durham JW (1994) **Fossil *Encope* (Echinoidea) from the Pacific coast of southern Mexico** *Revista Mexicana de Ciencias Geológicas* 11(1):113-116.

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

133. Ali MSM (1983) **The paleogeographic distribution of *Clypeaster* (Echinoidea) during the Cenozoic Era** *Neues Jahrb für Geologie und Paläontologie Monatshefte* 8:449-464.

134. Via L, Padreny J (1970) **Dos nuevas especies de *Clypeaster* del Eoceno de Cataluña**. *Publicaciones del Instuto de Investigaciones Geológicas de la Diputación Provincial* 24:89-118.

135. Wade BS, Pearson PN, Berggren WA, Pälike H (2011) **Review and revision of Cenozoic tropical planktonic foraminiferal biostratigraphy and calibration to the geomagnetic polarity and astronomical time scale** *Earth-Science Reviews* 104:111-142.

136. Gold DP, Fenton JP, Casas-Gallego M, Novak V, Pérez-Rodríguez I, Cetean C, et al (2018) **The biostratigraphic record of Cretaceous to Paleogene tectono-eustatic relative sea-level change in Jamaica** *Journal of South American Earth Sciences* 86:140-161.

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

138. Reich A, Dunn C, Akasaka K, Wessel G (2015) **Phylogenomic analyses of Echinodermata support the sister groups of Asterozoa and Echinozoa** *PLoS One* 10:e0119627.

139. Linchangco Jr GV, Foltz DW, Reid R, Williams J, Nodzak C, Kerr AM, et al (2017) **The phylogeny of extant starfish (Asteroidea: Echinodermata) including Xyloplax, based on comparative transcriptomics** *Molecular Phylogenetics and Evolution* 115:161-170.

140. Janies DA, Witter Z, Linchangco GV, Foltz DW, Miller AK, Kerr AM, et al (2016) **EchinoDB, an application for comparative transcriptomics of deeply-sampled clades of echinoderms** *BMC Bioinformatics* 17:17-48.

141. Clouse RM, Linchangco Jr GV, Kerr AM, Reid RW, Janies DA (2015) **Phylotranscriptomic analysis uncovers a wealth of tissue inhibitor of metalloproteinases variants in echinoderms** *Royal Society Open Science* 2:150377.

1245 142. Koga H, Fujitani H, Morino Y, Miyamoto N, Tsuchimoto J, Shibata TF, et al (2016) **Experimental**
1246 **approach reveals the role of alx1 in the evolution of the echinoderm larval skeleton** *PLoS One*
1247 11:e0149067.
1248 143. Simakov O, Kawashima T, Marlétaz F, Jenkins J, Koyanagi R, Mitros T, et al (2015) **Hemichordate**
1249 **genomes and deuterostome origins** *Nature* 527:459-465.
1250 144. Collins M, Peck LS, Clark MS (2021) **Large within, and between, species differences in marine**
1251 **cellular responses: Unpredictability in a changing environment** *Science of The Total Environment*
1252 148594.