

1 **Phylogenomic analyses shed light on the relationships of chiton superfamilies and**
2 **shell-eye evolution**

3

4 Xu Liu^{1,2}, Julia D. Sigwart^{3*}, Jin Sun^{1,2*}

5

6 ¹ Institute of Evolution and Marine Biodiversity, Ocean University of China, Qingdao 266003,
7 China

8 ² Laoshan Laboratory, Qingdao, China

9 ³ Department of Marine Zoology, Senckenberg Research Institute and Natural History
10 Museum Frankfurt, 60325 Frankfurt am Main, Germany

11

12 * Corresponding author, Jin Sun: jin_sun@ouc.edu.cn or Julia D. Sigwart:

13 julia.sigwart@senckenberg.de

14

15 **Abstract**

16 Mollusca is the second-largest animal phylum with over 100,000 species among eight distinct
17 taxonomic classes. Across 1000 living species in the class Polyplacophora, chitons have a
18 relatively constrained morphology but with some notable deviations. Several genera possess
19 “shell eyes”, true eyes with a lens and retina that are embedded within the dorsal shells,
20 which represent the most recent evolution of animal eyes. The phylogeny of major chiton
21 clades is mostly well established, in a set of superfamily and higher-level taxa supported by
22 various approaches including multiple gene markers, mitogenome-phylogeny and
23 phylotranscriptomic approaches as well as morphological studies. However, one critical
24 lineage has remained unclear: *Schizochiton* was controversially suggested as a potential
25 independent origin of chiton shell eyes. Here, with the draft genome sequencing of
26 *Schizochiton incisus* (superfamily Schizochitonoidea) plus assembly of transcriptome data
27 from other polyplacophorans, we present phylogenetic reconstructions using both
28 mitochondrial genomes and phylogenomic approaches with multiple methods. Phylogenetic
29 trees from mitogenomic data are inconsistent, reflecting larger scale confounding factors in
30 molluscan mitogenomes. A consistent robust topology was generated with protein coding
31 genes using different models and methods. Our results support Schizochitonoidea is a sister
32 group to other Chitonoidea in Chitonina, in agreement with established classification. This
33 suggests that the earliest origin of shell eyes is in Schizochitonoidea, which were also gained
34 secondarily in other genera in Chitonoidea. Our results have generated a holistic review of
35 the internal relationship within Polyplacophora, and a better understanding on the evolution
36 of Polyplacophora.

37

38 **Keywords:** Polyplacophora, Chiton, Phylogenomics, Mollusca, shell eyes

39 **Introduction**

40 Molluscs represent the second most species rich animal phylum with the broadest
41 morphological disparity of body plans. The class Polyplacophora, also known as chitons,
42 includes around 1000 living species and over 400 fossil species (Stebbins et al. 2009).
43 Chitons are exclusively marine, and their most distinctive feature is eight separate aragonitic
44 valves or plates on their dorsal side (Ladd 1966; Stebbins et al. 2009; Irisarri et al. 2020).
45 They attach to the substratum with a muscular ventral foot and feed with an iron-mineralised
46 radula (Joester et al. 2016). They have no head or cephalised senses, and therefore lack
47 conventional eyes. However, the dorsal valves are densely innervated with a complex array of
48 sensory pores called aesthetes which can have densities of over 1000 mm⁻².

49
50 Aesthete pores are present in all chitons, with substantial differences in morphology, size,
51 arrangement, densities, and presumably also functions, and aesthete morphology is often used
52 to discriminate species in taxonomic descriptions (Sirenko 2006). A number of chiton species
53 are demonstrably photosensitive, and some have pigmented aesthetes that apparently function
54 as photoreceptors. In the most elaborate variation, in a few genera, some of the larger
55 “megalaesthete” pores have further developed into shell eyes. These are true eyes, embedded
56 in the shell matrix, with a crystalline lens and a pigmented photoreceptive retina (Sigwart et
57 al. 2021).

58
59 The evolution of chiton shell eyes occurred much more recently than any other animal eyes.
60 The oldest fossil shell eyes are known from the fossil genus *Incissiochiton* from the lower
61 Palaeocene (61-66 Mya), which is a member of the family Schizochitonidae, the only family
62 in a superfamily Schizochitonoidea (Sirenko 2006; Sirenko 2013). Members of
63 Schizochitonidae (*Incissiochiton* and the Recent genus *Schizochiton*), as well as species in the
64 two subfamilies Acanthopleurinae and Toniciinae, possess shell eyes. The only previous
65 molecular phylogenetic study that included *Schizochiton* dates back to 2003 with five gene
66 fragments (Okusu et al. 2003), and those authors suggested that the phylogenetic position of *S.*
67 *incisus* in those analyses was “unstable” and deserved further discussion. Most importantly,
68 the unresolved phylogenetic position of *Schizochiton* raised the possibility that shell eye
69 structures evolved not only relatively recently, but in two separate events. However, in the
70 last 20 years this hypothesis has not been tested further, due to a lack of appropriate specimen
71 material for molecular data from this important lineage *Schizochiton*.

72
73 Phylogenetic systematics of Polyplacophora has been developed using both morphological
74 and molecular characters (Albano 2021). Extant chitons are divided into three well-resolved
75 orders: Lepidopleurida, Callochitonida, and Chitonida (Giribet et al. 2020). Lepidopleurida
76 consists of mainly deep-sea species with distinctive morphological synapomorphies including

77 aesthete arrangement, gills, and a specialized sense organ called the Schwabe Organ (Sigwart
78 et al. 2014). The position of *Callochiton* was equivocal in earlier studies but usually resolved
79 as sister to Chitonida (Koch et al. 1990; Sigwart et al. 2013) and the single family
80 Callochitonidae is now recognized as comprising a separate order-ranked clade
81 Callochitonida (Sigwart et al. 2013; Giribet et al. 2020; Moles et al. 2021). Most living
82 chitons are in the order Chitonida, which is further divided into two suborders, Chitonina
83 (including two superfamilies, Chitonoidea and Schizochitonoidea) and Acanthochitonina
84 (including two superfamilies, Mopaliaoidea and Cryptoplacoidea). The backbone phylogeny
85 of chitons is well understood especially at the level of superfamilies, for all clades except for
86 Schizochitonoidea.

87

88 Various genomic and transcriptomic data in Polyplacophora are now available on NCBI but
89 were generated independently for several different research purposes (Table 1). There are
90 only two chiton genomes available, *Acanthopleura granulata* (Varney et al. 2021) and
91 *Hanleya hanleyi* (Varney et al. 2022). Meanwhile, two independent phylogenomic studies
92 based on transcriptome sequencing, generated data for species and genera that cover all
93 Recent superfamilies: *Callochiton*, *Tonicia schrammi*, *Chiton tuberculatus*, *Chiton*
94 *marmoratus*, *Chaetopleura apiculata*, *Lepidozonia mertensii*, *Mopalia muscosa*, *Katharina*
95 *tunicata*, *Tonicella lineata*, *Nutallochiton* sp., *Cryptoplax japonica* and *Cryptoplax*
96 *larvaeformis* (Varney et al. 2021) and *Lepidopleurus cajetanus* (SRX5063921), *Callochiton*
97 *septemvalvis*, *Stenoplax bahamensis*, *Cryptoplax japonica* and *Choneplax lata* (Moles et al.
98 2021). There are also some other studies examining the gene expression profiles, which
99 includes *Leptochiton cascadiensis* (Halanych et al. 2014), *Acanthopleura lochooana* (Liu et
100 al. 2022), *Rhyssoplax olivacea* (Riesgo et al. 2012), *Cryptochiton stelleri* (Nemoto et al.
101 2019), *Acanthochitona crinita* (De Oliveira et al. 2016), *Acanthochitona rubrolineata*
102 (SRP179406) and *Acanthochitona fascicularis* (SRR13862580). These data collection can
103 support a phylogenomic construction with larger taxon coverage. And due to the important
104 position of *Schizochiton* for us to better understand chiton evolution, we newly sequenced and
105 assembled the genome and mitogenome of *Schizochiton incisus*. Combining this with other
106 available chiton data from NCBI and previous studies, we aimed to reconstruct a phylogeny
107 of Polyplacophora at the superfamily level with different phylogenomics inferences and tree
108 reconstruction methods, specifically to test the position of *S. incisus* and Schizochitonoidea.

109

110 **Material and Methods**

111 **1. Sample collection**

112 All genomes and transcriptomes used in this study are listed in Table 1. To increase taxon
113 sampling, we newly sequenced an individual of *Schizochiton incisus* and also *Leptochiton*
114 *asellus*. *Schizochiton incisus* was collected from a rock on a coral reef at the depth of 80 m of

115 Livock Reef (10°10'N, 115°19'E) by fishing net in the South China Sea on July 11, 2020 (Fig.
116 1). The whole animals *S. incisus* was preserved in 95% EtOH, which was later stored at room
117 temperature, and a small piece of girdle tissue was removed for DNA extraction. The *S.*
118 *incisus* sample was deposited in the malacology collections at the Senckenberg Museum,
119 Frankfurt with catalogue number SMF 386201. *Leptochiton asellus* was collected on the
120 rocky shore in September 2019, at Ballyhenry Island, Strangford Lough, at Portaferry, N.
121 Ireland. For *L. asellus*, five tissues, including foot, perinotum, aesthetes, viscera, and shell
122 edge were dissected and fixed in RNAlater (ThermoFisher) at 4 degree and transferred to -80
123 deep freezer for storage.

124

125 **2. Genome and RNA sequencing**

126 Total genomic DNA of *S. incisus* was extracted with a DNeasy Blood & Tissue Kit (QIAGEN,
127 Germantown, Maryland), which was further sequenced for 150bp paired-end Illumina
128 sequencing to generate approximately 40Gb of raw data on NovaSeq 6000 platform at
129 Novogene (Beijing).

130

131 RNA of *Leptochiton asellus* was extracted using Trizol (ThermoFisher) and sent to Novogene
132 (Beijing) for Eukaryotic type transcriptome library preparation and further sequenced on
133 NovaSeq 6000 platform. Approximately 6Gb of raw reads were generated for each tissue.

134

135 **3. Mitogenome analysis**

136 **3.1 Mitogenome assemble and annotation**

137 The raw data were trimmed using Trimmomatic v.0.39 (Bolger et al. 2014) with strict
138 filtering settings (ILLUMINACLIP: adapters.fa:2:30:10 LEADING:20 TRAILING:20
139 SLIDINGWINDOW:4:20 MINLEN:140) to remove low-quality reads and adapters
140 contaminated reads. The resultant clean reads were initially assembled by SPAdes v.3.15.3
141 (Prjibelski et al. 2020) with default settings, and then the partial COI sequence of *S. incisus*
142 was extracted from the assembled contigs, which was later used as the “seed input” in
143 NOVOplasty v.4.2 (Dierckxsens et al. 2016) to obtain the complete mitogenome of *S. incisus*.
144 The mitogenome was then annotated using the MITOS web server (Donath et al. 2019) with
145 the invertebrate genetic code and the rest default settings, followed by a manual mitogenome
146 annotation confirmation by comparing with other chiton mitogenomes (Irisarri et al. 2020).

147

148 **3.2 Matrix construction**

149 All .gb files of chiton mitogenomes available on NCBI were downloaded and imported into
150 Phylosuite v.1.2.2 (Zhang et al. 2020), which is an application that allows users to perform
151 phylogenetic analyses on relatively small datasets. All procedures of mitogenome
152 phylogenetic analyses, except for tree constructing and visualization, were carried out

153 through Phylosuite built-in plugins. In brief, 13 protein-coding genes and 2 rRNA genes were
154 extracted from the chiton mitogenomes. Afterwards, MAFFT v. 7.471 was used to align
155 sequences, followed by trimAL v. 1.2rev57 with the “automated1” option to remove spurious
156 sequences and misaligned regions. After that, trimmed sequences were concatenated,
157 generating 3 different matrices. Amino acid sequences of 13 protein-coding genes (PCGs)
158 were extracted and concatenated into a Matrix1. As for Matrix2, all nucleotides of 13 PCGs
159 and 2 rRNA were concatenated. To avoid the phylogenetic signal saturation on the third
160 codon, the third codons of 13 PCGs were replaced by degenerate bases (A, G replaced by R
161 and C, T replaced by Y), then these modified sequences were concatenated, named Matrix3.
162 Generated gene matrix and the corresponding partition file were later used for maximum
163 likelihood (ML) and Bayesian inference (BI) tree construction.

164

165 **3.3 Mitogenome phylogeny**

166 For the ML framework, IQ-Tree v.2.1.3 (Minh et al. 2020) was implemented using -MFP to
167 select the best-fit model for each partition. Besides, an additional empirical profile mixture
168 model, C60, was also carried out on the AA matrix (Matrix1). All ML analysis were
169 performed with 1000 replicates of ultrafast bootstrapping (-bb 1000).

170

171 BI was carried out using PhyloBayes MPI v.1.8c (Lartillot et al. 2013) with CAT-GTR+ Γ 4
172 models. For each matrix, four independent Monte Carlo Markov chains (MCMC) were run
173 simultaneously and convergence was checked with the bpcomp program. Then a consensus
174 tree was obtained after discarding the first 10% cycles as a burn-in.

175 All trees obtained were then visualized with Figtree
176 (<http://tree.bio.ed.ac.uk/software/figtree/>).

177

178 **4. Genome assembly and annotation**

179 The Illumina raw data was filtered with Trimmomatic v.0.39 (Bolger et al. 2014) with
180 settings of “PE ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:10 TRAILING:10
181 SLIDINGWINDOW:4:15 MINLEN:40”. Afterwards, genome features were calculated by
182 using jellyfish v.2.3.0 (Marçais et al. 2011) (19mer) and GenomeScope2 (Vurture et al. 2017).
183 A benchmark of commonly used assemblers for Illumina data, including Platanus v.1.2.4
184 (Kajitani et al. 2014) and MaSuRCA v.4.0.3 (Zimin et al. 2013), was performed based on
185 BUSCO v.5.1.2 score by searching against metazoan odb10 database. Afterwards, purge-dups
186 v.1.2.5 (Guan et al. 2020) was used to remove redundant contigs, and the resultant contigs
187 were further scaffolded by using PEP-scaffolder (Zhu et al. 2016) with the help of protein
188 sequences from the concatenation of the genome of *Acanthopleura granulata* (Varney et al.
189 2021).

190

191 A custom repeat library of *S. incisus* was *de novo* generated by RepeatModeler v.2.0.2a
192 (Flynn et al. 2020). RepeatMasker v.4.1.0 (Tarailo-Graovac et al. 2009) was performed with
193 the species-specific repeat library mentioned above, followed by a second round of
194 RepeatMasker but with Repbase library 2018 (<https://www.girinst.org/replib/>). Afterwards,
195 BRAKER v.2.1.6 (Hoff et al. 2019) was run to train an *ab initio* gene predictor Augustus
196 v.3.4.0 (Stanke et al. 2006) with ODB10 v.1 database downloaded from OrthoDB
197 (Kriventseva et al. 2018), generating a config file of *S. incisus*, which was used as one piece
198 of evidence while running the genome annotator MAKER v.3.01.04 (Holt et al. 2011).
199 Because there was no transcript evidence available, all Mollusca proteins on NCBI were
200 downloaded (Date: Jan 20 2022), and redundancy was removed with CD-HIT v.4.8.1 with the
201 setting of “-c 0.9”. These protein sequences were regarded as the protein homology evidence
202 in MAKER. And the proteins generated from MAKER was used for further phylogenetic
203 analyses.

204

205 **5. Transcriptome assembly and filtration**

206 The protein coding genes of *Acanthopleura granulata*, *A. loochooana* (Liu et al. 2022) and
207 all other available transcriptomes were downloaded from NCBI SRA database. For
208 transcriptome SRA datasets as well as the transcriptome sequencing of *L. asellus*, the raw
209 reads were *de novo* assembled in Trinity v.2.13.2 or v.2.14.0 (Haas et al. 2013), using the
210 “--trimmomatic” setting, followed by one round of CD-HIT v.4.8.1 (Fu et al. 2012) with the
211 strictest threshold (-c 0.8) to remove redundant sequences. CD-HIT was run multiple times
212 which was continuously monitored by BUSCO5 aiming to get a best score with highest “S”
213 score and lowest “D” (duplicated BUSCO) score. Afterwards, Transdecoder v.5.5.0 (Douglas
214 2018) was performed to search for open reading frames with the “--single_best_only” option.
215 And the generated peptide files were filtered using CD-HIT with the “-c 0.8” option again to
216 make sure the “D” score wouldn’t drop any more. This step aimed to remove as many
217 heterozygous and transcript isoforms as possible so that they would not mislead orthology
218 inference.

219

220 **6. Orthology inference and matrix construction**

221 Orthology inference was accomplished with a pipeline that was generated from former
222 studies (Kocot et al. 2017; Sun et al. 2021) with slight modifications. We ran Orthofinder
223 v.2.5.4 (Emms et al. 2019) to search for orthologues within selected taxa. Then in the
224 “Orthogroup_Sequences” directory of the Orthofinder output, OG heads were fixed with a
225 custom shell script to make sure that the orthology inference pipeline could be error less.
226 After the preparation, PREQUAL v.1.02 (Whelan et al. 2018) was used to detect and mask
227 non-homologous characters. Then sequences shorter than 100 amino acids were deleted.
228 Occupancy was set to 50%, and redundant sequences were then removed with another custom

229 shell script named `uniqHaplo.pl`. The leftover `.fasta` files were aligned using MAFFT v.7.490
230 (Kato et al. 2013) with default settings. Afterwards, HmmCleaner (Di Franco et al. 2019)
231 was used to remove misaligned regions, followed by trimming alignment with BMGE v.1.12
232 (Crisuolo et al. 2010). Then FastTree2 (Price et al. 2010) was used to construct fast-ML
233 trees for each remaining OGs. Last but not least, PhyloPyPruner v.1.2.4
234 (<https://pypi.org/project/phylo-pypruner>) was performed to identify putative orthology
235 sequences based on the former FastTree2 result, resulting in an initial matrix containing 3593
236 OGs.

237

238 We performed genesortR (Mongiardino Koch 2021) to sort and select “best” OGs based on
239 seven commonly used phylogenetic gene properties, thus genes with best phylogenetic
240 signals can be used for down streaming analysis. An ML tree for the initial matrix was
241 constructed with the IQ-Tree “-MFP” model as input. Also, ML trees for each gene were
242 constructed in IQ-Tree with the same settings. At last, four matrices, including an initial
243 matrix (Matrix1), best 800 genes matrix (Matrix2), best 1300 genes matrix (Matrix3), and
244 best 2700 genes matrix (Matrix4) generated by genesortR, were prepared for phylogenetic
245 analysis.

246

247 **7. Phylogenomics**

248 ML phylogenetic analysis was performed using IQ-Tree 2 (Minh et al. 2020) on the four
249 matrices generated above. The ML approach was carried out using the best-fitting model for
250 each partition (-m MFP). Regarding the `.contree` file generated by the MFP model as the
251 guide tree, PMSF model was then performed in IQ-Tree 2 with site-specific frequency
252 models (C20, C40 and C60). All ML analyses were carried out with 1000 ultrafast bootstrap.
253 As for BI analysis, all matrices mentioned above were too large to run in PhyloBayes MPI
254 v.1.8c, thus the fifth matrix, produced by random 300 genes from Matrix1, was brought out.
255 Four independent chains were run simultaneously until convergent with CAT-GTR+Γ4
256 model.

257 A coalescent approach, in contrast to concatenated-based phylogenetic analysis, was also
258 performed to evaluate evolutionary relationships in polyplacophora with ASTRAL v.5.7.1
259 (Sayyari et al. 2016). An AU-test was performed with IQ-tree 2 on two topologies, which
260 were ((Chitonoidea, Schizochitonoidea), Acanthochitonina) and ((Acanthochitonina,
261 Schizochitonoidea), Chitonoidea), respectively.

262

263 **Results**

264 **Mitochondrial genome**

265 We assembled the complete mitochondrial genome of *S. incisus*, which was 15,491 bp in
266 length circularized with 13 PCGs, 2 rRNA, and 22 tRNA, a typical mitogenome architecture

267 of bilaterians. Protein-coding genes are coded with normal invertebrate mitochondrial codons
268 including the start and stop codons. The mitogenome of *S. incisus* follows the proposed
269 hypothetical ancestral gene order for Polyplacophora (Irisarri et al. 2020), except for an
270 inversion of trnG-trnE (Fig. 2b). The mitogenome gene order seems to be relatively
271 conserved in Polyplacophora compared to those in gastropods or bivalves (Irisarri et al.
272 2020).

273

274 **Mitochondrial phylogeny**

275 The phylogenetic trees reconstructed with mitogenome data showed significant discordance
276 among different methods and matrices. There were 3 distinct topologies for the position of *S.*
277 *incisus*, which were ((Chitonoidea, Schizochitonoidea), Acanthochitonina)(13PCGs with
278 MFP, PB based on modified 3rd codon), ((Acanthochitonina, Schizochitonoidea),
279 Chitonoidea) (13PCGs with C60, PB, PCGs + rRNA with MFP, PCGs + rRNA with PB) and
280 ((Chitonoidea, Acanthochitonina), Schizochitonoidea) (modified 3rd codon), respectively.
281 The statistical support of the *S. incisus* node was lower than 95% in all methods, except for
282 BI, indicating these nodes were not well supported with mitogenomic data. We note that in
283 addition to *Schizochiton*, the position of *Plaxiphora albida* also varied from one clade to
284 another (Fig. 2a). And in the presentative tree, *Tonicina zschau*i was sister to the rest
285 Chitonoidea.

286

287 **Genome and transcriptome assembly**

288 Genome features of *S. incisus* were estimated with Illumina sequencing reads, which resulted
289 in an estimated genome size of 1.1 GB and genome heterozygosity of 0.93%. Draft genome
290 assembly from MaSuRCA generated a better result (C:73.8% [S:68.1%, D:5.7%]) than the
291 Platanus version [C:17.9% (S:13.7%, D:4.2%)], which was used for down-stream analyses.
292 After further scaffolding with protein sequences from other chitons with available genomes
293 and removing heterozygous contigs, the final assembly has a BUSCO score of C:73.8%, N50
294 of 13.2Kb and the assembled size of 971 Mb.

295

296 By collecting the evidence from the *ab initial* method and protein evidence, a total of 23,444
297 protein coding genes were predicted in *S. incisus* with a BUSCO score of C: 40.8% (S: 37.0%,
298 D: 3.8%) and F: 19.8%. Though the score is lower than the *Acanthochitona rubrolineata*
299 genome (Varney et al. 2021), 12,419 of them (52%) can find their reciprocal best hits BLAST
300 in *A. rubrolineata*, suggesting that a good coverage of protein coding genes for the
301 phylogenomic analyses.

302

303 The transcriptome of *Leptochiton asellus* generated from five tissues was assembled into
304 390,724 contigs with an N50 value of 1.68Kb, and the BUSCO score is C:94.5%

305 (S:83.1%,D:11.4%). For the rest transcriptome assembly of the publicly available data, the
306 BUSCO completeness ranges from 12.9% (*Lepidopleurus cajetanus*) to 95.8% (*Callochiton*
307 *septemvalvis*) (for the species list and their corresponding BUSCO score, see Table 1).

308

309 **Phylogenomics**

310 The phylogenomic analysis was based on the combination of transcriptome and genome data,
311 covering all the extant superfamilies in Polyplacophora (Table 1). There were four matrices
312 generated by genesortR forming seven distinct phylogenetic signals. Minimum occupancy for
313 all matrices was set to 50%. The sites contained in the four matrices are 696,897 (3593 genes,
314 all genes, Matrix 1), 194,356 (best 800 genes, Matrix 2), 299,710 (best 1300 genes, Matrix 3),
315 554,857 (best 2700 genes, Matrix 4), respectively (Fig. 3).

316

317 The phylogenetic trees reconstructed from nuclear data, including coalescent approach results,
318 showed a high degree of consistency about the position of *S. incisus*, as sister to Chitonoidea
319 (Fig. 4). Support for this Schizochitonoidea + Chitonoidea clade retrieved node support of
320 100% in all analyses except for PMSF-C20 of Matrix1 (which is 59), showing a relatively
321 stable topology. The support for all superfamily level groups and their arrangement was
322 consistently high. However, the positions of some tips are unsettled: *Chaetopleura apiculata*,
323 *Lepidozonia mertensii* and *Stenoplax bahamensis* resolved in variable positions within the
324 superfamilies. The relationship of *Choneplax* relative to the members of genus
325 *Acanthochitona* is also changeable.

326

327 **Topology test**

328 We performed AU-test on two topologies based on Matrix1 to determine the better supported
329 tree topology. Given results with P -value < 0.05 will be rejected. The results showed that the
330 first tree topology ((Chitonoidea, Schizochitonoidea), Acanthochitonina) was accept with a
331 P -value of 0.952, and the second topology ((Acanthochitonina, Schizochitonoidea),
332 Chitonoidea) was rejected with a P -value of 0.0476.

333

334 **Discussions**

335 The phylogenetic relationships of chiton at the order and superfamily levels are relatively
336 stable and well resolved. Based on a consensus of phylogenetic analyses, Polyplacophora is
337 divided into three orders, Lepidopleurida, Callochitonida and Chitonida (Irisarri et al. 2020;
338 Moles et al. 2021), which is also recovered in the present analyses. At the superfamily level,
339 former molecular studies lacked data to test the position of Schizochitonoidea, and our results
340 support the sister relationship of Chitonoidea + Schizochitonoidea in a monophyletic
341 suborder Chitonina, as proposed from integrated morphological and anatomical evidence
342 (Sirenko 2006).

343

344 The mitogenome data were much less informative than nuclear transcriptome and genomic
345 data. We used mitogenome data of available chitons to reconstruct phylogenetic trees with
346 different approaches, including ML and Bayesian inference, but the results below superfamily
347 level are unstable. For example, in the representative tree selected for mitochondrial analyses,
348 *Tonicina zschau* formed a sister group to other remaining Chitonoidea, whereas current
349 systematics would predict a placement for *Tonicina* within the small clade formed by the
350 genera *Lepidozona*, *Ischnochiton* and *Chaetopleura*. The topology we illustrated is not
351 supported by 4 of 7 trees reconstructed by corresponding methods, so this placement should
352 be taken as unresolved. As already suggested in the previous mitogenome phylogeny of
353 chitons (Irisarri et al. 2020), this could be a result by poor taxon sampling, but it was not
354 improved by adding a few additional taxa here. Indeed, this issue of low phylogenetic signal
355 in mitogenome phylogeny has also been raised in data from another molluscan class,
356 Monoplacophora (Stoger et al. 2016), and confounding features occur in many molluscan
357 mitogenomes (Ghiselli et al. 2021).

358

359 Interestingly, *Schizochiton* possesses a unique mitogenome gene order, differing from any
360 other chitons with available mitogenomes, which might imply relatively fast evolution of the
361 species. Mitogenome phylogenies are currently not reliable for reconstructing detailed
362 phylogenies for Polyplacophora and potentially other molluscan clades. This may be
363 improved with better taxon sampling, or may be a fundamental problem of insufficient
364 phylogenetic signal. It is clear that currently phylogenomic approaches are needed to
365 reconstruct phylogeny of chitons at or below superfamily level resolution.

366

367 All the phylogenomic results for the main lineage in this study shared the same topology with
368 strong node support except for Matrix1-C20. The topology is consistent with what is by now
369 a well-established backbone phylogeny for Polyplacophora and also concordant at
370 superfamily and higher level with the mitogenome phylogeny (Sigwart et al. 2013; Irisarri et
371 al. 2020; Moles et al. 2021). Lepidopleurida is sister to the remaining Polyplacophora.
372 *Callochiton*, representing the order Callochitonida is sister to Chitonida. This latter order is
373 divided into two clear clades representing the suborders Chitonina and Acanthochitonina.

374

375 Our phylogeny of polyplacophora based on phylogenomic approach possesses more
376 advantages than former molecular studies (Okusu et al. 2003; Sigwart et al. 2013; Irisarri et al.
377 2020; Moles et al. 2021) , including a broader taxon sampling and massive genes, has
378 resolved the relationships among main lineages of chitons. The genus and family level
379 arrangement of taxa in this study are largely concordant with established taxonomy or with
380 other molecular studies from smaller data matrices. Within Lepidopleurida, the family

381 Leptochitonidae s.s. is restricted to the NE Atlantic species, represented here by *Leptochiton*
382 *asellus* and *Lepidopleurus cajetaus*, with the Pacific *Leptochiton cascadiensis* outside that
383 clade, as is already established from previous molecular studies using Sanger sequencing
384 (Sigwart et al. 2011; Sigwart 2016).

385

386 Acanthochitonina is known to be divided into two clades based on egg hulls and hexagon
387 edges projections; one of the clades, Mopaliioidea, includes *Cryptochiton*, *Mopalia*,
388 *Katharina* and *Tonicella* (Okusu et al. 2003) which is also well supported by every other
389 molecular phylogeny including our results and modern phylogenetic systematics (Sigwart et
390 al. 2013). Family level arrangement is difficult to test with limited taxon sampling, but the
391 genera in our study group into these four genera that are closely allied to Mopaliidae as
392 separate from a second clade of *Nuttallina* + *Cyanoplax*, also as found in previous studies
393 (Irisarri et al. 2020). In the other superfamily Cryptoplacoidea, *Nuttallochiton* is sister to the
394 rest of Cryptoplacoidea, in accordance with previous molecular studies (Okusu et al. 2003;
395 Sigwart et al. 2013; Irisarri et al. 2020). However, the position of *Plaxiphora* within
396 Acanthochitonina is equivocal; this has been a persistent problem in every molecular
397 phylogeny of chitons, although multiple morphological characters unite *Plaxiphora* with the
398 family Mopaliidae (Sirenko 2006).

399

400 *Schizochiton* resolved as sister to Chitonoidea, forming a monophyletic suborder Chitonina
401 with full support except for Matrix1-C20 method, being sister group with the larger order
402 Chitonida. The only prior molecular analysis to include *Schizochiton* also recovered it as
403 sister to the remaining Chitonina in one version of their analyses, but concluded that its
404 position within the phylogeny was effectively unresolved (Okusu et al. 2003: fig 5). The
405 position of *Schizochiton* was controversial because of an unusual combination of
406 morphological characters. The balance of evidence placed this group in the suborder
407 Chitonina (Sirenko 2006). *Schizochiton* possess a caudal sinus in tail valve that is similar to
408 others in Mopaliioidea as well as egg hulls with cupules that are simpler but comparable to
409 other Mopaliioidea. Based on the new phylogenetic tree, we can infer these features may be
410 plesiomorphic for the larger order Chitonida.

411

412 One important morphological feature of *Schizochiton* that differs from almost all other
413 chitons is their shell eyes. Shell eyes were described in 1884 from specimens of *Schizochiton*
414 *incissus* (Moseley 1884), and were immediately recognized as modifications of the chiton
415 aesthete system (Moseley 1885). All chitons possess aesthete pores in their shell plates and
416 some are photosensitive (Kingston et al. 2018). But shells eyes are restricted to only a few
417 genera, in the family Schizochitonidae and the family Chitonidae. Those genera in the family
418 Chitonidae with shell eyes form a monophyletic clade and have a fossil record only dating

419 back to the Miocene (Sirenko 2006). Phylogenetic and fossil evidence suggests that shell eyes
420 evolved first in Schizochitonidae and again a second time very recently in the history of
421 Chitonidae.

422

423 Recognizing *Schizochiton* within a superfamily level group Schizochitonoidea, sister to
424 Chitonoidea, confirms the relationship predicted by morphological systematics. This is now
425 confirmed from molecular evidence and a more stable phylogeny than earlier preliminary
426 results. This also reaffirms that the multiple lines of evidence from morphological,
427 anatomical, and gamete characters already recognized in chitons provide a robust basis for
428 phylogenetic systematics.

429

430 **Data availability**

431 The raw Illumina sequencing data was deposited on NCBI SRA database with the accession
432 No. of PRJNA909482, and the assembled mitogenome on NCBI nucleotide database with the
433 No. of XXXX. The assembled genomic contigs, predicted gene models can be accessed via
434 FigShare with the URL of [10.6084/m9.figshare.21709742](https://www.figshare.com/figure/21709742).

435

436 **Acknowledgements**

437 This research project was financially supported by the Fundamental Research Funds for the
438 Central Universities (202241002 and 202172002), Science and Technology Innovation
439 Project of Laoshan Laboratory (No. LSKJ202203100), and the Young Taishan Scholars
440 Program of Shandong Province (tsqn202103036). Bioinformatic analysis was conducted on
441 the high-performance server IEMB-1 hosted at Institute of Evolution and Marine Biodiversity.
442 We also thank Dr. Chong Chen (JAMSTEC) for help with fieldwork and specimens. This is
443 contribution number 14 from the Senckenberg Ocean Species Alliance.

444

445 **References**

- 446 Albano PG (2021). *Biology and evolution of the Mollusca*, WILEY 111 RIVER ST, HOBOKEN
447 07030-5774, NJ USA.
- 448 Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data.
449 *Bioinformatics* (Oxford, England) 30(15): 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
- 450 Criscuolo A, Gribaldo S (2010). BMGE (Block Mapping and Gathering with Entropy): a new
451 software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC*
452 *Evol Biol* 10: 210. <https://doi.org/10.1186/1471-2148-10-210>
- 453 De Oliveira A, Wollesen T, Kristof A, Scherholz M, Redl E, Todt C, Bleidorn C, Wanninger A (2016).
454 Comparative transcriptomics enlarges the toolkit of known developmental genes in mollusks. *BMC*
455 *genomics* 17(1): 1-23.
- 456 Di Franco A, Poujol R, Baurain D, Philippe H (2019). Evaluating the usefulness of alignment filtering
457 methods to reduce the impact of errors on evolutionary inferences. *BMC Evol Biol* 19(1): 21.
458 <https://doi.org/10.1186/s12862-019-1350-2>
- 459 Dierckxsens N, Mardulyn P, Smits G (2016). NOVOPlasty: de novo assembly of organelle genomes
460 from whole genome data. *Nucleic Acids Research* 45(4): e18-e18. <https://doi.org/10.1093/nar/gkw955>
- 461 Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, Middendorf M, Bernt M
462 (2019). Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes.
463 *Nucleic Acids Research* 47(20): 10543-10552. <https://doi.org/10.1093/nar/gkz833>
- 464 Douglas (2018). TransDecoder/TransDecoder. GitHub. Available from:
465 <https://github.com/TransDecoder/TransDecoder> (accessed March 23, 2020).
- 466 Emms DM, Kelly S (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics.
467 *Genome Biol* 20(1): 238. <https://doi.org/10.1186/s13059-019-1832-y>
- 468 Flynn JM, Hubble R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF (2020). RepeatModeler2
469 for automated genomic discovery of transposable element families. *Proceedings of the National*
470 *Academy of Sciences* 117(17): 9451-9457. <https://doi.org/doi:10.1073/pnas.1921046117>
- 471 Fu L, Niu B, Zhu Z, Wu S, Li W (2012). CD-HIT: accelerated for clustering the next-generation
472 sequencing data. *Bioinformatics* 28(23): 3150-3152. <https://doi.org/10.1093/bioinformatics/bts565>
- 473 Ghiselli F, Gomes-Dos-Santos A, Adema CM, Lopes-Lima M, Sharbrough J, Boore JL (2021).
474 Molluscan mitochondrial genomes break the rules. *Philos Trans R Soc Lond B Biol Sci* 376(1825):
475 20200159. <https://doi.org/10.1098/rstb.2020.0159>
- 476 Giribet G, Edgecombe GD (2020). *The Invertebrate Tree of Life*, Princeton University Press.
- 477 Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R (2020). Identifying and removing
478 haplotypic duplication in primary genome assemblies. *Bioinformatics* 36(9): 2896-2898.
479 <https://doi.org/10.1093/bioinformatics/btaa025>
- 480 Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B,
481 Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T,
482 Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A (2013). De novo transcript sequence
483 reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature*
484 *Protocols* 8(8): 1494-1512. <https://doi.org/10.1038/nprot.2013.084>
- 485 Halanych KM, Kocot KM (2014). Repurposed transcriptomic data facilitate discovery of innate
486 immunity toll-like receptor (TLR) genes across Lophotrochozoa. *The Biological Bulletin* 227(2):
487 201-209.
- 488 Hoff KJ, Lomsadze A, Borodovsky M, Stanke M (2019). *Whole-Genome Annotation with BRAKER*.
489 *Methods in molecular biology* (Clifton, N.J.) 1962: 65-95.
490 https://doi.org/10.1007/978-1-4939-9173-0_5
- 491 Holt C, Yandell M (2011). MAKER2: an annotation pipeline and genome-database management tool
492 for second-generation genome projects. *BMC Bioinformatics* 12(1): 491.
493 <https://doi.org/10.1186/1471-2105-12-491>
- 494 Irisarri I, Uribe JE, Eernisse DJ, Zardoya R (2020). A mitogenomic phylogeny of chitons (Mollusca:
495 Polyplacophora). *BMC Evol Biol* 20(1): 22. <https://doi.org/10.1186/s12862-019-1573-2>
- 496 Joester D, Brooker LR (2016). The Chiton *Radula*: A Model System for Versatile Use of Iron Oxides*.
497 *Iron Oxides*: 177-206.
- 498 Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu
499 E, Maruyama H, Kohara Y, Fujiyama A, Hayashi T, Itoh T (2014). Efficient de novo assembly of
500 highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res* 24(8):

- 501 1384-1395. <https://doi.org/10.1101/gr.170720.113>
- 502 Katoh K, Standley DM (2013). MAFFT multiple sequence alignment software version 7:
503 improvements in performance and usability. *Mol Biol Evol* 30(4): 772-780.
504 <https://doi.org/10.1093/molbev/mst010>
- 505 Kingston ACN, Chappell DR, Speiser DI (2018). Evidence for spatial vision in *Chiton tuberculatus*, a
506 chiton with eyespots. *J Exp Biol* 221(Pt 19). <https://doi.org/10.1242/jeb.183632>
- 507 Koch RA, Lambert CC (1990). Ultrastructure of sperm, spermiogenesis, and sperm-egg interactions
508 in selected invertebrates and lower vertebrates which use external fertilization. *J Electron Microsc*
509 *Tech* 16(2): 115-154. <https://doi.org/10.1002/jemt.1060160204>
- 510 Kocot KM, Struck TH, Merkel J, Waits DS, Todt C, Brannock PM, Weese DA, Cannon JT, Moroz LL,
511 Lieb B, Halanych KM (2017). Phylogenomics of Lophotrochozoa with Consideration of Systematic
512 Error. *Syst Biol* 66(2): 256-282. <https://doi.org/10.1093/sysbio/syw079>
- 513 Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM (2018).
514 OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for
515 evolutionary and functional annotations of orthologs. *Nucleic Acids Research* 47(D1): D807-D811.
516 <https://doi.org/10.1093/nar/gky1053>
- 517 Ladd HS (1966). Chitons and gastropods (Haliotidae through Adeorbidae) from the western Pacific
518 Islands. *Professional Paper*.
- 519 Lartillot N, Rodrigue N, Stubbs D, Richer J (2013). PhyloBayes MPI: Phylogenetic Reconstruction
520 with Infinite Mixtures of Profiles in a Parallel Environment. *Systematic Biology* 62(4): 611-615.
521 <https://doi.org/10.1093/sysbio/syt022>
- 522 Liu C, Liu H, Huang J, Ji X (2022). Optimized Sensory Units Integrated in the Chiton Shell. *Marine*
523 *Biotechnology*. <https://doi.org/10.1007/s10126-022-10114-2>
- 524 Marçais G, Kingsford C (2011). A fast, lock-free approach for efficient parallel counting of
525 occurrences of k-mers. *Bioinformatics* 27(6): 764-770. <https://doi.org/10.1093/bioinformatics/btr011>
- 526 Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R
527 (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic
528 Era. *Molecular Biology and Evolution* 37(5): 1530-1534. <https://doi.org/10.1093/molbev/msaa015>
- 529 Moles J, Cunha TJ, Lemer S, Combosch DJ, Giribet G (2021). Tightening the girdle:
530 phylotranscriptomics of Polyplacophora. *Journal of Molluscan Studies* 87(2).
531 <https://doi.org/10.1093/mollus/eyab019>
- 532 Mongiardino Koch N (2021). Phylogenomic Subsampling and the Search for Phylogenetically
533 Reliable Loci. *Mol Biol Evol* 38(9): 4025-4038. <https://doi.org/10.1093/molbev/msab151>
- 534 Moseley HN (1884). XIX.—On the presence of eyes and other sense-organs in the shells of the
535 Chitonidae. *The Annals and magazine of natural history; zoology, botany, and geology* 14: 141-147.
536 <https://doi.org/10.1080/00222938409459782>
- 537 Moseley HN (1885). *Memoirs: on the presence of eyes in the shells of certain Chitonidæ, and on the*
538 *structure of these organs. Journal of Cell Science* 2(97): 37-60.
- 539 Nemoto M, Ren D, Herrera S, Pan S, Tamura T, Inagaki K, Kisailus D (2019). Integrated
540 transcriptomic and proteomic analyses of a molecular mechanism of radular teeth biomineralization in
541 *Cryptochiton stelleri*. *Scientific reports* 9(1): 1-10.
- 542 Okusu A, Schwabe E, Eernisse DJ, Giribet G (2003). Towards a phylogeny of chitons (Mollusca,
543 Polyplacophora) based on combined analysis of five molecular loci. *Organisms Diversity & Evolution*
544 3(4): 281-302. <https://doi.org/10.1078/1439-6092-00085>
- 545 Price MN, Dehal PS, Arkin AP (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for
546 Large Alignments. *PLOS ONE* 5(3): e9490. <https://doi.org/10.1371/journal.pone.0009490>
- 547 Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A (2020). Using SPAdes De Novo
548 Assembler. *Current Protocols in Bioinformatics* 70(1): e102.
549 <https://doi.org/https://doi.org/10.1002/cpbi.102>
- 550 Riesgo A, Andrade S, Sharma PP, Novo M, Pérez-Porro AR, Vahtera V, González VL, Kawauchi GY,
551 Giribet G (2012). Comparative description of ten transcriptomes of newly sequenced invertebrates
552 and efficiency estimation of genomic sampling in non-model taxa. *Frontiers in zoology* 9(1): 1-24.
- 553 Sayyari E, Mirarab S (2016). Fast Coalescent-Based Computation of Local Branch Support from
554 Quartet Frequencies. *Molecular Biology and Evolution* 33(7): 1654-1668.
555 <https://doi.org/10.1093/molbev/msw079>
- 556 Sigwart JD (2016). Deep trees: Woodfall biodiversity dynamics in present and past oceans. *Deep Sea*
557 *Research Part II: Topical Studies in Oceanography* 137: 282-287.

- 558 <https://doi.org/10.1016/j.dsr2.2016.06.021>
- 559 Sigwart JD, Schwabe E, Saito H, Samadi S, Giribet G (2011). Evolution in the deep sea: a combined
560 analysis of the earliest diverging living chitons (Mollusca: Polyplacophora: Lepidopleurida).
561 *Invertebrate Systematics* 24(6): 560-572. <https://doi.org/10.1071/IS10028>
- 562 Sigwart JD, Stoeger I, Knebelberger T, Schwabe E (2013). Chiton phylogeny (Mollusca :
563 Polyplacophora) and the placement of the enigmatic species *Chorioplax grayi* (H. Adams & Angas).
564 *Invertebrate Systematics* 27(6): 603. <https://doi.org/10.1071/is13013>
- 565 Sigwart JD, Sumner-Rooney L (2021). Continuous and Regular Expansion of a Distributed Visual
566 System in the Eyed Chiton *Tonicia lebruni*. *Biol Bull* 240(1): 23-33. <https://doi.org/10.1086/712114>
- 567 Sigwart JD, Sumner-Rooney LH, Schwabe E, Heß M, Brennan GP, Schrödl M (2014). A new sensory
568 organ in “primitive” molluscs (Polyplacophora: Lepidopleurida), and its context in the nervous system
569 of chitons. *Frontiers in Zoology* 11(1): 7. <https://doi.org/10.1186/1742-9994-11-7>
- 570 Sirenko B (2006). New Outlook on the System of Chitons (Mollusca: Polyplacophora) (the 2nd
571 International Chiton Symposium). *Venus (Journal of the Malacological Society of Japan)* 65: 27-49.
572 https://doi.org/10.18941/venus.65.1-2_27
- 573 Sirenko B (2013). Four new species and one new genus of Jurassic chitons (Mollusca: Polyplacophora:
574 Lepidopleurida) from the Middle Russian Sea. *Proc. Zool. Inst. RAS* 317: 30-44.
575 <https://doi.org/10.31610/trudyzin/2013.317.1.30>
- 576 Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B (2006). AUGUSTUS: ab initio
577 prediction of alternative transcripts. *Nucleic Acids Research* 34(suppl_2): W435-W439.
578 <https://doi.org/10.1093/nar/gk1200>
- 579 Stebbins TD, Eernisse DJ (2009). Chitons (Mollusca: Polyplacophora) known from benthic
580 monitoring programs in the Southern California Bight. *The Festivus* 41(6): 53-100.
- 581 Stoger I, Kocot KM, Poustka AJ, Wilson NG, Ivanov D, Halanych KM, Schrod M (2016).
582 Monoplacophoran mitochondrial genomes: convergent gene arrangements and little phylogenetic
583 signal. *BMC Evol Biol* 16(1): 274. <https://doi.org/10.1186/s12862-016-0829-3>
- 584 Sun J, Li R, Chen C, Sigwart JD, Kocot KM (2021). Benchmarking Oxford Nanopore read
585 assemblers for high-quality molluscan genomes. *Philos Trans R Soc Lond B Biol Sci* 376(1825):
586 20200160. <https://doi.org/10.1098/rstb.2020.0160>
- 587 Tarailo-Graovac M, Chen N (2009). Using RepeatMasker to identify repetitive elements in genomic
588 sequences. *Curr Protoc Bioinformatics* Chapter 4: Unit 4.10.
589 <https://doi.org/10.1002/0471250953.bi0410s25>
- 590 Varney RM, Speiser DI, McDougall C, Degnan BM, Kocot KM (2021). The Iron-Responsive
591 Genome of the Chiton *Acanthopleura granulata*. *Genome Biol Evol* 13(1).
592 <https://doi.org/10.1093/gbe/evaa263>
- 593 Varney RM, Yap-Chiongco MK, Mikkelsen NT, Kocot KM (2022). Genome of the lepidopleurid
594 chiton *Hanleya hanleyi* (Mollusca, Polyplacophora). *F1000Research* 11(555): 555.
- 595 Vurture GW, Sedlazeck FJ, Nattestad M, Underwood CJ, Fang H, Gurtowski J, Schatz MC (2017).
596 GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* 33(14):
597 2202-2204. <https://doi.org/10.1093/bioinformatics/btx153>
- 598 Whelan S, Irisarri I, Burki F (2018). PREQUAL: detecting non-homologous characters in sets of
599 unaligned homologous sequences. *Bioinformatics* 34(22): 3929-3930.
600 <https://doi.org/10.1093/bioinformatics/bty448>
- 601 Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT (2020). PhyloSuite: An integrated and
602 scalable desktop platform for streamlined molecular sequence data management and evolutionary
603 phylogenetics studies. *Molecular Ecology Resources* 20(1): 348-355.
604 <https://doi.org/10.1111/1755-0998.13096>
- 605 Zhu BH, Song YN, Xue W, Xu GC, Xiao J, Sun MY, Sun XW, Li JT (2016). PEP_scaffolder: using
606 (homologous) proteins to scaffold genomes. *Bioinformatics* 32(20): 3193-3195.
607 <https://doi.org/10.1093/bioinformatics/btw378>
- 608 Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA (2013). The MaSuRCA genome
609 assembler. *Bioinformatics* 29(21): 2669-2677. <https://doi.org/10.1093/bioinformatics/btt476>
- 610
- 611

612 **Figure legend:**

613 Figure 1. *Schizochiton incisus* (b) and the position where it was collected (a, marked with a
614 red spot). C shows the whole genomic pipeline used in this study, including sample
615 preparation, mitogenome analysis, draft genome assembling and annotation and
616 phylogenomic approach. Photos courtesy of Prof. Xiaoqi Zheng.

617

618 Figure 2. Mitogenome analyses of *Schizochiton incisus*; (a) mitogenome phylogeny of
619 Polyplacophora, and (b) *S. incisus* mitochondrial gene order comparing with hypothetical
620 ancestral mitochondrial gene order of chitons.

621

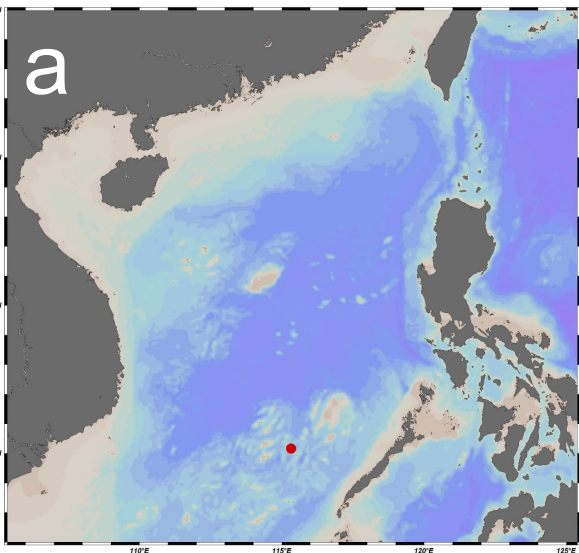
622 Figure 3. The occupancy of the four matrices generated by genesortR.

623

624 Figure 4. Phylogeny of chiton based on phylogenomic approach with different methods.
625 Node support are transferred into matrices colored with a continuous scale bar ranging from 0
626 to 1. Blue indicates 100% support and pink indicates the topology is not supported by the
627 representative tree. And node with blue spot indicates full support in all methods. M1-M5,
628 matrix 1-5; MFP, IQ-Tree MFP model; C20-C60, profile mixture models C20-C60; M1 Astral,
629 coalescent analysis based on Matrix1; M5 PB, PhyloBayes analysis based on Matrix5.

Table 1 Statistics of chiton genomes and transcriptomes used in this study, including number of contigs and BUSCO scores after filtering.

Species	SRA No.	No. of proteins	BUSCO score	Source
Lepidopleurida				
<i>Hanleya hanleyi</i>	SRR11674123	47,786	C:81.7%[S:80.7%,D:1.0%]	Varney, Yap-Chiongco et al. 2022
<i>Lepidopleurus cajetanus</i>	SRX5063921	10,479	C:12.9%[S:12.7%,D:0.2%]	-----
<i>Leptochiton asellus</i>	-----	81,610	C:94.7%[S:86.9%,D:7.8%]	this study
<i>Leptochiton rugatus</i>	SRR1611558	23,030	C:79.4%[S:77.9%,D:1.5%]	Halanych and Kocot 2014
Callochitonida				
<i>Callochiton septemvalvis</i>	SRR13010089	30,618	C:95.9%[S:87.8%,D:8.1%]	Moles, Cunha et al. 2021
<i>Callochiton</i> sp.	SRR11674125	8235	C:28.2%[S:26.1%,D:2.1%]	Varney, Speiser et al. 2021
Chitonida				
<i>Acanthopleura granulata</i>	-----	19,621	C:93.8%[S:93.3%,D:0.5%]	Varney, Speiser et al. 2021
<i>Acanthopleura loochooana</i>	-----	44,182	C:90.4%[S:85.3%,D:5.1%]	Liu, Liu et al. 2022
<i>Tonicia schrammi</i>	SRR11674132	16,274	C:67.4%[S:67.1%,D:0.3%]	Varney, Speiser et al. 2021
<i>Chiton tuberculatus</i>	SRR11674134	18,002	C:83.2%[S:82.8%,D:0.4%]	Varney, Speiser et al. 2021
<i>Chiton marmoratus</i>	SRR11674135	5848	C:26.5%[S:26.5%,D:0.0%]	Varney, Speiser et al. 2021
<i>Rhysoplax olivacea</i>	SRR618506	27,356	C:67.1%[S:65.3%,D:1.8%]	Riesgo, Andrade et al. 2012
<i>Chaetopleura apiculata</i>	SRR11674124	18,915	C:79.3%[S:79.1%,D:0.2%]	Varney, Speiser et al. 2021
<i>Lepidozona mertensii</i>	SRR11674130	13,531	C:72.1%[S:71.5%,D:0.6%]	Varney, Speiser et al. 2021
<i>Stenoplax bahamensis</i>	SRR13010087	24,602	C:39.7%[S:39.0%,D:0.7%]	Moles, Cunha et al. 2021
<i>Schizochiton incisus</i>	-----	20,902	C:40.9%[S:37.5%,D:3.4%]	this study
<i>Cryptochiton stelleri</i>	DRP005555	19,101	C:82.2%[S:81.7%,D:0.5%]	Nemoto, Ren et al. 2019
<i>Mopalia muscosa</i>	SRR11577121	13,262	C:77.0%[S:76.6%,D:0.4%]	Varney, Speiser et al. 2021
<i>Katharina tunicata</i>	SRR11674131	15,542	C:89.7%[S:88.4%,D:1.3%]	Varney, Speiser et al. 2021
<i>Tonicella lineata</i>	SRR11577222	13,780	C:79.0%[S:77.7%,D:1.3%]	Varney, Speiser et al. 2021
<i>Nutallochiton</i> sp.	SRR11674133	57,110	C:74.3%[S:67.4%,D:6.9%]	Varney, Speiser et al. 2021
<i>Cryptoplax japonica</i>	SRR13010086	14,963	C:34.6%[S:34.3%,D:0.3%]	Moles, Cunha et al. 2021
<i>Cryptoplax larvaeformis</i>	SRR11674126	20,128	C:88.1%[S:87.7%,D:0.4%]	Varney, Speiser et al. 2021
<i>Choneplax lata</i>	SRR13010088	16,971	C:14.3%[S:13.4%,D:0.9%]	Moles, Cunha et al. 2021
<i>Acanthochitona rubrolineata</i>	SRP179406	44,221	C:91.8%[S:71.2%,D:20.6%]	-----
<i>Acanthochitona crinita</i>	SRR5110525	22,678	C:91.4%[S:91.0%,D:0.4%]	De Oliveira, Wollesen et al. 2016
<i>Acanthochitona fascicularis</i>	SRR13862580	17,427	C:88.9%[S:88.5%,D:0.4%]	-----



Other chitons' genomes & transcriptomes

Trinity

CD-HIT

Transdecoder

CD-HIT-est

"clean" data

MaSuRCA

purge-dups

PEP-scaffolder

Repeat Mask

BRAKER

MAKER

Fixed by ETOH

DNA extraction

Trimmomatic

Illumina sequencing

SPAdes

NOVOplasty

MITOS

Other chitons' mitogenomes

MAFFT

TrimAl

Matrix1

Matrix2

Matrix3

IQ-Tree C60

IQ-Tree MFP

PhyloBayes

C

Orthofinder

Delete short alignments

uniqHaplo.pl

MAFFT

Matrix1

PhyloPyPruner

FastTree2

BMGE

HmmCleaner

genesortR

Random 300 genes

Matrix1

Matrix2

Matrix3

Matrix4

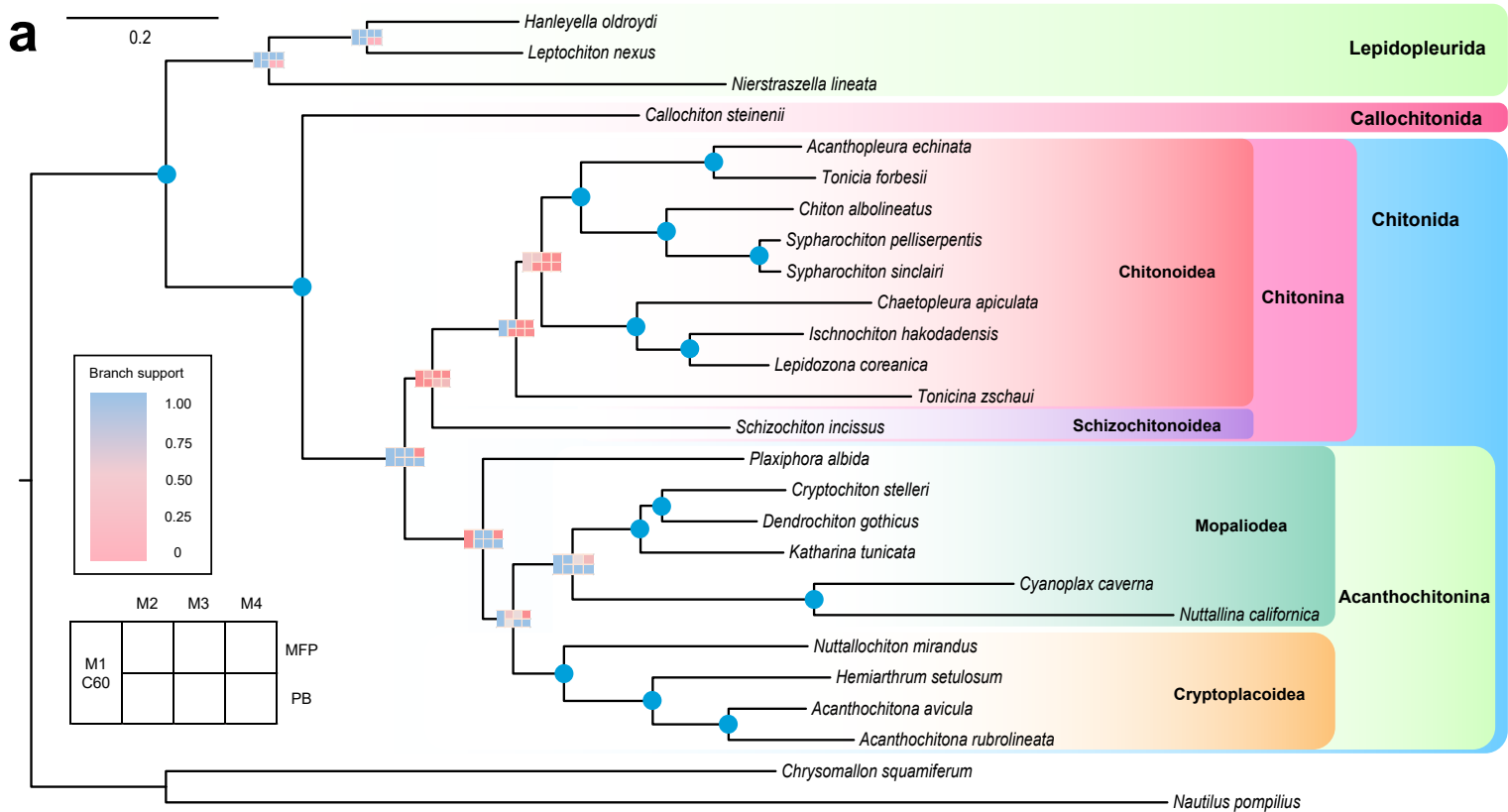
Matrix5

Astral

MCMCTree

IQTree MFP, C20, C40 & C60

PhyloBayes



b

Schizochiton incisus mitochondrial gene order

cox1 | cox2 | D | atp8 | atp6 | T | P | E | G | cox3 | K | A | R | N | I | nad3 | S1 | nad2
F | nad5 | H | nad4 | nad4L | S2 | cob | nad6 | nad1 | L2 | L1 | rrL | V | rrS | M | C | Y | W | Q

Hypothetical ancestral mitochondrial gene order of chitons

cox1 | cox2 | D | atp8 | atp6 | T | P | cox3 | K | A | R | N | I | nad3 | S1 | nad2
F | nad5 | H | nad4 | nad4L | S2 | cob | nad6 | nad1 | L2 | L1 | rrL | V | rrS | M | C | Y | W | Q | G | E

aligned length: 0-500

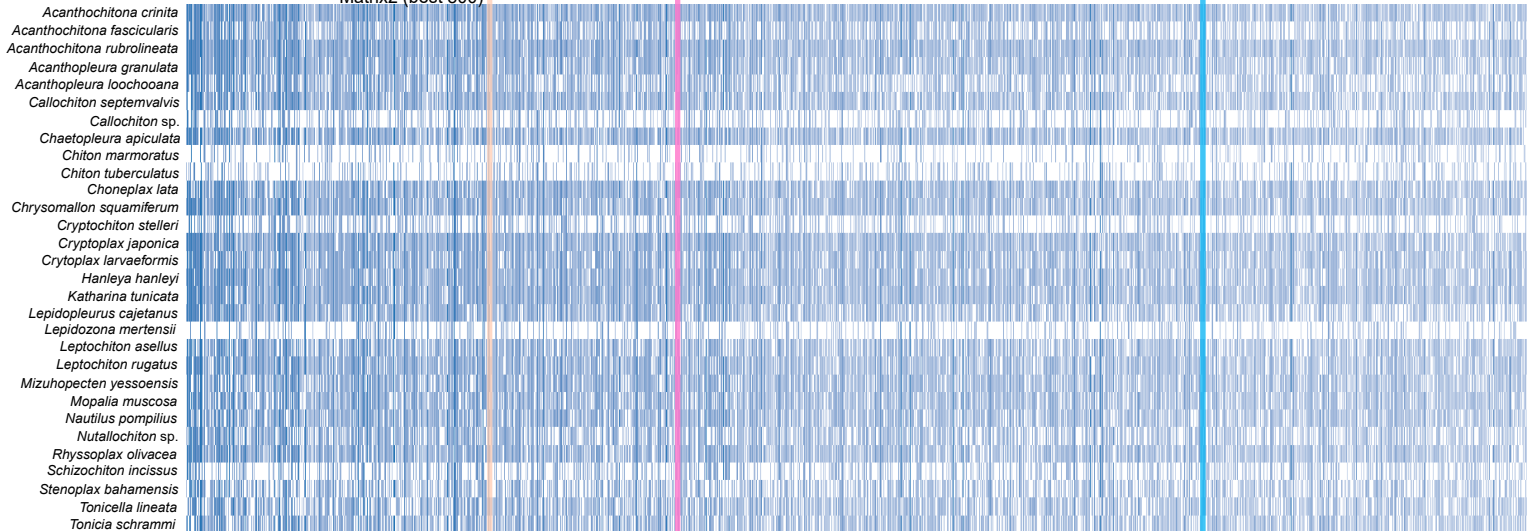


Matrix1 (all 3593)

Matrix4 (best 2700)

Matrix3 (best 1300)

Matrix2 (best 800)



← OGs sorted by genesrotR score

