



**HAL**  
open science

## Mitigating Anticipated Effects of Systematic Errors Supports Sister-Group Relationship between Xenacoelomorpha and Ambulacraria

Hervé Philippe, Albert J Poustka, Marta Chiodin, Katharina J Hoff,  
Christophe Dessimoz, Bartłomiej Tomiczek, Philipp H Schiffer, Steven Müller,  
Daryl Domman, Matthias Horn, et al.

► **To cite this version:**

Hervé Philippe, Albert J Poustka, Marta Chiodin, Katharina J Hoff, Christophe Dessimoz, et al.. Mitigating Anticipated Effects of Systematic Errors Supports Sister-Group Relationship between Xenacoelomorpha and Ambulacraria. *Current Biology - CB*, 2019, 29 (11), pp.1818-1826.e6. 10.1016/j.cub.2019.04.009 . hal-02392672

**HAL Id: hal-02392672**

**<https://hal.science/hal-02392672>**

Submitted on 17 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Mitigating anticipated effects of systematic errors supports sister-group relationship between Xenacoelomorpha and Ambulacraria.

Hervé Philippe<sup>1,2&</sup>, Albert J. Poustka<sup>3,4,&</sup>, Marta Chiodin<sup>5,6</sup>, Katharina J. Hoff<sup>7</sup>, Christophe Dessimoz<sup>8,9,10,11</sup>, Bartłomiej Tomiczek<sup>8,12</sup>, Philipp H. Schiffer<sup>8</sup>, Steven Müller<sup>8</sup>, Daryl Domman<sup>13,14</sup>, Matthias Horn<sup>13</sup>, Heiner Kuhl<sup>15,16</sup>, Bernd Timmermann<sup>15</sup>, Noriyuki Satoh<sup>17</sup>, Tomoe Hikosaka-Katayama<sup>18</sup>, Hiroaki Nakano<sup>19</sup>, Matthew L. Rowe<sup>20</sup>, Maurice R. Elphick<sup>20</sup>, Morgane Thomas-Chollier<sup>21</sup>, Thomas Hankeln<sup>22</sup>, Florian Mertes<sup>23</sup>, Andreas Wallberg<sup>24</sup>, Jonathan P. Rast<sup>25</sup>, Richard R. Copley<sup>26</sup>, Pedro Martinez<sup>4,27</sup>, Maximilian J. Telford<sup>8\*</sup>

## Affiliations:

1 Centre de Théorisation et de Modélisation de la Biodiversité, Station d'Ecologie Théorique et Expérimentale, UMR CNRS 5321, Moulis 09200, France.

2 Département de Biochimie, Centre Robert-Cedergren, Université de Montréal, Montréal, QC, Canada H3C 3J7.

3 Max-Planck Institute for Molecular Genetics, Evolution and Development Group, Ihnestrasse 73, Berlin 14195, Germany and Dahlem Centre for Genome Research

4 Medical Systems Biology, Environmental and Phylogenomics Group, Max-Planck-Straße 3, 12489 Berlin, Germany

5 Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Av. Diagonal, 643, Barcelona 08028, Spain

6 New York University, School of Medicine, 435 E 30th St, New York, NY 10016

7 Bioinformatics Group, Institute for Mathematics and Computer Science, University of Greifswald, Walther-Rathenau-Str. 47, 17487 Greifswald, Germany

8 Centre for Life's Origins and Evolution, Department of Genetics, Evolution and Environment, University College London, Darwin Building, Gower Street, London, WC1E 6BT

Department of Computer Science, University College London, Darwin Building, Gower Street, London, WC1E 6BT, UK

9 Department of Computational Biology, University of Lausanne, 1015 Lausanne, Switzerland

10 Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland

31 11 Swiss Institute of Bioinformatics, Génopode, 1015 Lausanne, Switzerland  
32 12 Laboratory of Evolutionary Biochemistry, Intercollegiate Faculty of Biotechnology, University  
33 of Gdansk and Medical University of Gdansk, Gdansk, 80-307, Poland  
34 13 Department of Microbiology and Environmental Systems Science, University of Vienna, A-  
35 1090 Vienna, Austria  
36 14 present address: Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM  
37 87545, USA  
38 15 Sequencing Core Facility, Max Planck Institute for Molecular Genetics, Ihnestraße 63-73  
39 14195 Berlin, Germany  
40 16 Department of Ecophysiology and Aquaculture, Leibniz Institute for Freshwater Ecology and  
41 Inland Fisheries, Müggelseedamm 301, 12587 Berlin, Germany.  
42 17 Marine Genomics Unit, Okinawa Institute of Science and Technology Graduate University,  
43 Onna, Okinawa 904-0495, Japan  
44 18 Natural Science Center for Basic Research and Development, Gene Science Division,  
45 Hiroshima University, Higashi-Hiroshima 739-8527, Japan  
46 19 Shimoda Marine Research Center, University of Tsukuba, Shimoda, Shizuoka 415-0025,  
47 Japan  
48 20 Queen Mary University of London, School of Biological & Chemical Sciences, Mile End  
49 Road, London, E1 4NS, UK  
50 21 Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Ecole Normale Supérieure,  
51 CNRS, INSERM, PSL Université Paris, 75005 Paris, France.  
52 22 Johannes Gutenberg University Mainz, Institute of Organismic and Molecular Evolution,  
53 Molecular Genetics and Genome Analysis, J. J. Becher-Weg 30a, 55128 Mainz, Germany  
54 23 Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt  
55 (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg, Gemany  
56 24 Department of Medical Biochemistry and Microbiology, Science for Life Laboratory, Uppsala  
57 University, SE -751 23, Uppsala, Sweden.  
58 25 Department of Pathology and Laboratory Medicine, Emory University School of Medicine,  
59 Atlanta, GA, USA

60 26 Sorbonne Université, CNRS, Laboratoire de Biologie du Développement de Villefranche-sur-  
61 mer (LBDV), 06230 Villefranche-sur-mer, France

62 27 ICREA (Institut Català de Recerca i Estudis Avancats), Passeig Lluís Companys 23  
63 08010 Barcelona, Spain

64

65

66 & These authors contributed equally to this work

67 \* Correspondence: [m.telford@ucl.ac.uk](mailto:m.telford@ucl.ac.uk)

68

69 Lead Contact: Max Telford

70

71 Keywords: *Xenoturbella*, Acoelomorpha, Ambulacraria, Nephrozoa, phylogenomics,  
72 systematic error, phylogeny, Metazoa

73

#### 74 **Summary.**

75 *Xenoturbella* and the acoelomorph worms (Xenacoelomorpha) are simple marine animals with  
76 controversial affinities. They have been placed as the sister group of all other bilaterian  
77 animals (Nephrozoa hypothesis) implying their simplicity is an ancient characteristic [1, 2];  
78 alternatively, they have been linked to the complex Ambulacraria (echinoderms and  
79 hemichordates) in a clade called the Xenambulacraria [3-5], suggesting their simplicity  
80 evolved by reduction from a complex ancestor. The difficulty resolving this problem implies  
81 the phylogenetic signal supporting the correct solution is weak and affected by inadequate  
82 modelling, creating a misleading non-phylogenetic signal. The idea that the Nephrozoa  
83 hypothesis might be an artefact is prompted by the faster molecular evolutionary rate observed  
84 within the Acoelomorpha. Unequal rates of evolution are known to result in the systematic  
85 artefact of long branch attraction which would be predicted to result in an attraction between  
86 long branch acoelomorphs and the outgroup pulling them towards the root [6]. Other biases  
87 inadequately accommodated by the models used can also have strong effects, exacerbated

88 in the context of short internal branches and long terminal branches [7]. We have assembled  
89 a large and informative data set to address this problem. Analyses designed to reduce or to  
90 emphasise misleading signals show the Nephrozoa hypothesis is supported under conditions  
91 expected to exacerbate errors and the Xenambulacraria hypothesis is preferred in conditions  
92 designed to reduce errors. Our reanalyses of two other recently published data sets [1, 2]  
93 produce the same result. We conclude that the Xenacoelomorpha are simplified relatives of  
94 the Ambulacraria.

95

## 96 **Results**

### 97 *Assembling our data matrix*

98 In order to provide the best chance of avoiding artefacts generated by data errors [7, 8] we  
99 assembled a new data set of 1,173 genes (350,088 amino acid positions) from a balanced  
100 and rich selection of 59 taxa with just 23.5% missing data, giving us a matrix that is larger and  
101 more complete than any previously used to examine the question. Our new matrix has been  
102 carefully curated to minimise potential errors from sources including contamination and non-  
103 orthology. Alongside existing data, it includes new gene predictions from 6 partial genomes  
104 and 4 new transcriptomes.

105

106 New predicted protein sets were derived from partial genomes of *Xenoturbella bocki*,  
107 *Symsagittifera roscoffensis*, *Meara stichopi*, *Nemertoderma westbladi*, *Pseudaphanostoma*  
108 *variabilis* and *Praesagittifera naikaiensis*; from new transcriptomes of *Xenoturbella bocki*,  
109 *Symsagittifera roscoffensis*, *Paratomella rubra* and *Isodiametra pulchra* and from published  
110 data available at the NCBI. To produce a balanced and computationally tractable data set we  
111 selected approximately equal numbers (6-8) of diverse species from the following clades:  
112 Xenacoelomorpha, Hemichordata, Echinodermata, Chordata, Lophotrochozoa, Ecdysozoa,  
113 Cnidaria and Porifera plus the placozoan *Trichoplax adhaerens*. We omitted members of  
114 Ctenophora due to their well-documented fast evolutionary rate [9]. From these original sets

115 of predicted protein sequences, we used OMA to identify probable groups of orthologs  
116 covering the Metazoa [10, 11]. As OMA is rather stringent and can therefore omit valid  
117 orthologs, we added some missing orthologs using the 42 pipeline  
118 (<https://bitbucket.org/dbaurain/42/downloads/>). These putative orthologs were then tested for  
119 possible cross contamination, non-orthology and other issues likely to affect accurate  
120 phylogenetic reconstruction (see methods). Our final data set contained 1,173 orthologous  
121 genes from 59 species of animals giving a total of 350,088 aligned amino acids.

122

### 123 *Comparisons with existing recent data matrices*

124 We compared our matrix to the two most recent studies addressing the question of the  
125 affinities of the Xenacoelomorpha in terms of data quality (percent of clades present in the  
126 concatenated tree that are also present in single gene trees) and quantity (number of amino  
127 acids present in the supermatrix: this number comes from the total number of amino acids in  
128 the matrix; if there were no missing data this would equal length of alignment multiplied by the  
129 number of species). Our dataset is among the largest and of the highest quality: our single-  
130 gene trees recover >50% on average of the expected clades, whereas the average for the  
131 other data sets is 29% (maximum 39% - See Figure S4D). This indicates that our dataset likely  
132 contains fewer erroneous data (e.g. contaminants, paralogs, frameshifts) than others and is  
133 therefore likely to contain more genuine phylogenetic signal: a prerequisite to infer  
134 phylogenies accurately [7, 9].

135

### 136 *Analyses of our data using site heterogeneous models show limited support for* 137 *Xenambulacraria*

138 We analysed our complete matrix using a gene jackknife approach, which provides a  
139 conservative measure of clade support while being computationally tractable [9]. We used  
140 cross validation to compare the fit of different models of sequence evolution on all data sets  
141 and found that the CATGTR model was the best fitting in all cases. We therefore used the

142 CATGTR model of PhyloBayes [12] with a gamma correction for between site rate variability  
143 to analyse 100 subsamples each containing ~90,000 positions from the complete data set.  
144 We found weak support (60% jackknife support) for a monophyletic grouping of  
145 Xenacoelomorpha and Ambulacraria. The second best supported topology grouped  
146 Xenacoelomorpha with Protostomes (24% jackknife support) and Nephrozoa had 13%  
147 jackknife support. Other uncontroversial clades in the tree were reconstructed with strong  
148 support (Figure 1A,B). In common with some previously published results [13, 14], the  
149 relationships between Chordata, Xenambulacraria and Protostomia were unresolved - we did  
150 not reconstruct a monophyletic Deuterostomia (Chordata plus (Xen)ambulacraria).

151

#### 152 *Removing fast evolving Acoelomorpha reduces support for Nephrozoa*

153 Our approach to testing the possible effects of systematic error is to consider situations in  
154 which we can predict whether, if the tree is influenced by artefacts, nodal support will increase  
155 or decrease using different subsets of data or analytical methods. Manipulations expected to  
156 strengthen artefactual signal (less adequate models or subsets of data with an exaggerated  
157 systematic bias) are expected to increase support for the artefactual topology and vice versa,  
158 while the genuine phylogenetic signal should remain unaffected. One established approach  
159 for dealing with LBA is to remove the fastest evolving members of the group of interest [6]. If  
160 the Nephrozoa signal depends on an LBA artefact, we predict support for Nephrozoa would  
161 decrease in favour of Xenambulacraria when fast evolving members of Xenacoelomorpha are  
162 removed. The Acoelomorpha have clearly evolved more rapidly than *Xenoturbella* (Figure 1A)  
163 and this difference seems to be mirrored in the more derived gene content of acoelomorph  
164 genomes [15, 16].

165

166 The validity of this approach requires the Xenacoelomorpha to be monophyletic. In our  
167 jackknife tree, and in previous phylogenomic analyses, the Xenoturbellida is strongly  
168 supported as the sister group of Acoelomorpha. This conclusion is further supported by a

169 *Xenoturbella*/Acoelomorpha specific rare genomic change involving their Caudal/CDX  
170 ortholog (Figure S4E). If we therefore accept Xenacoelomorpha as monophyletic, it is  
171 legitimate to use the slowly evolving member of the clade (*Xenoturbella*) as a representative  
172 of the Xenacoelomorpha, so reducing the effects of rapid evolution in the Acoelomorpha.  
173 When we removed the long branched Acoelomorpha but included the slower evolving  
174 *Xenoturbella* and repeated the jackknifing of the complete data set, the support for  
175 Xenambulacraria increased to 81% (Figure 1C). This result is consistent with the support for  
176 Xenacoelomorpha being reduced in part due to LBA caused by the fast evolving  
177 Acoelomorpha.

178

179 *Stratifying genes according to phylogenetic accuracy: genes with difficult to extract*  
180 *phylogenetic signal support Nephrozoa*

181 A given gene is expected to vary in its ability to reconstruct the phylogeny of interest according  
182 to the method being used. More accurate genes ('better' genes with respect to the  
183 phylogenetic method used) will have more appropriate or more even rates of substitution or,  
184 more generally, some genes may fit the assumptions of the models used more closely than  
185 others; equally, some alignments may contain non-orthologous - e.g. contaminant -  
186 sequences. We reason that the genes that perform best at reconstructing known clades with  
187 a given method should be the most reliable when solving a related phylogenetic problem. To  
188 stratify the genes in our concatenated alignment according to their ability to reconstruct an  
189 accurate tree, we measured the capacity of each gene to reconstruct uncontroversial  
190 monophyletic groups of animals using two different methods that gave virtually identical  
191 results. After stratifying our genes, we concatenated them in order from best to worst and took  
192 the genes covering the first 25% of genes (best) and those covering the last 25% of genes  
193 (worst). The proportions of missing data and constant positions were similar for the two sub-  
194 datasets, but the worst genes evolved faster and were more saturated (Table 1); CATGTR is  
195 the best fitting model in each case, and improvement over GTR seems to be more important  
196 for the worst genes (Table 1). Posterior predictive checks show that the best genes violate the

197 models much less than the worst genes (Table 1), but that even the best fitting CATGTR  
198 model does not explain the data well. We performed gene jackknife analysis with CATGTR  
199 using 50 samples of ~30,000 positions. The best performing genes according to our criterion  
200 supported Xenambulacraria (including the long branched acoelomorphs) with 94% jackknife  
201 support (Figure 1D). The worst genes supported Nephrozoa with a weak 48% jackknife  
202 support and we observed lower support for other clades across the tree in agreement with the  
203 expected difficulty in extracting phylogenetic signal from these genes. The best genes also  
204 support Xenambulacraria (JP= 63%) when the short branched *Xenoturbella* is removed  
205 leaving just the fast evolving Acoelomorpha (Figure 2S). Since the genes with the better  
206 phylogenetic to non-phylogenetic signal ratio consistently support Xenambulacraria, the likely  
207 explanation is that support for Nephrozoa is an artefact caused by the limitations of  
208 reconstruction methods when applied to problematic data.

209

210 *Better fitting models support Xenambulacraria, worse models support Xenambulacraria if long*  
211 *branch acoelomorphs are removed*

212 Consistent with previous studies [5, 17, 18], the site heterogeneous CATGTR model we used  
213 has a better fit to our data set than the site homogeneous LG and GTR models predominantly  
214 used by Cannon et al. and Rouse et al. [1, 2] (cross-validation score of  $3034 \pm 152$  and  $2001$   
215  $\pm 155$ , respectively). While we have shown the best genes analysed with CATGTR support  
216 Xenambulacraria even with long branch Acoelomorpha included, analysing this data set with  
217 less well-fitting site homogeneous GTR models supports Nephrozoa (100% bootstrap  
218 support). When reanalysing the best data after removing the long branched Acoelomorpha,  
219 however, even the less well fitting GTR model supports Xenambulacraria (92% bootstrap  
220 support, Figure 2). For the worst performing genes, all analyses (CATGTR, and GTR with or  
221 without Acoelomorpha) supported Nephrozoa (Figure 2). Data and analyses that are better by  
222 specified, measurable, objective criteria consistently result in increased support for  
223 Xenambulacraria.

224

225 *Addressing the effects of compositional bias reduces support for Nephrozoa*

226 After LBA, probably the best-known source of systematic error is compositional bias, in which  
227 a systematic tendency of substitutions towards certain amino acids in subsets of taxa affects  
228 tree reconstruction [19]. Considering the possibility that compositional biases in the  
229 proportions of amino acids found in different species were inadequately accounted for by the  
230 models used, we looked for evidence of the existence of compositional bias by using posterior  
231 predictive checks in PhyloBayes to compare real amino acid frequencies of the 59 species in  
232 our data with their mean values under the null distribution predicted by the best fitting CATGTR  
233 model. A strong compositional bias was observed in our data although not specifically in  
234 Xenacoelomorpha. Interestingly, part of the superiority of the 'better' genes discussed  
235 previously may be explained by the lower compositional bias we observe in the best 25% of  
236 data compared to the worst 25% (mean squared heterogeneity - best genes = ~100; worst  
237 genes = ~190). If compositional bias is contributing to the support for Nephrozoa, then  
238 reducing the effects of this bias would be predicted to lower support for Nephrozoa. To  
239 minimise the effects of species specific compositional bias we recoded the amino acids in our  
240 alignment using a reduced alphabet that gathers similar (and frequently substituted) amino  
241 acids into the following 6 'Dayhoff' groups (A,G,P,S,T) (D,E,N,Q) (H,K,R) (F,Y,W) (I,L,M,V)  
242 (C). Recoding also tends to reduce model violations and saturation as frequently substituting  
243 amino acids are consolidated into a single character state [19]. We reran the jackknife  
244 analyses of the complete data set using the recoded data in PhyloBayes [12]. Using all species  
245 and all genes, jackknife support for Xenambulacraria increased from 61% to 90% suggesting  
246 that compositional bias affects tree reconstruction and specifically reduces support for  
247 Xenambulacraria (Figure 3). We repeated this analysis using a bootstrapping approach  
248 instead of jackknifing and the support for Xenambulacraria was found to be 98%. This increase  
249 is in line with other evidence indicating the relatively conservative nature of jackknife support  
250 values.

251

252 *The effects of model and data testing are not data set specific*

253 One possible criticism of our findings is that they depend on the particular subset of genes  
254 and taxa used. We repeated our analyses using the data sets of Cannon et al. and Rouse et  
255 al. [1, 2]. For each test (removing long branched taxa, stratifying genes according to  
256 phylogenetic accuracy and recoding to reduce compositional bias) we observed the same  
257 direction of change as we observe in our data, albeit with lower support values, especially for  
258 the taxon-poor Rouse et al data (see Figures S4). While Cannon et al. analysed their data  
259 with long branched Acoelomorpha omitted, they used the site homogenous LG model and  
260 recovered the Nephrozoa tree. Using CATGTR on the same data we recovered the  
261 Xenambulacraria tree (Figure S2). With the same results coming from three, large,  
262 independently assembled data sets it is reasonable to conclude that the support for  
263 Xenambulacraria cannot be explained by the choices made during data set assembly.

264

## 265 **Discussion**

266 Determining the correct phylogenetic position of the Xenacoelomorpha has significant  
267 implications for our understanding of their evolution and that of the Metazoa. If  
268 Xenacoelomorpha diverged prior to other bilaterian animals, then this could explain their  
269 relative morphological simplicity and lack, for example, of several bilaterian Hox genes and  
270 microRNAs [20-22]. Under the assumption of such an 'early-diverging' scenario,  
271 xenacoelomorphs were naturally considered to be of particular interest, as a branch  
272 intermediate between non-bilaterians (such as Cnidaria) and Nephrozoa [23, 24]. If, on the  
273 other hand, xenacoelomorphs are the sister group of the Ambulacraria, their simplicity, both  
274 morphological and genetic, must have been derived from a more complex ancestor by a  
275 process of character loss. If we accept that the Xenambulacraria clade is real, we should  
276 expect additional evidence for this relationship to remain in the embryology, morphology and  
277 genomes of these animals and such evidence would be a valuable corroboration of our results.

278 Although it seems that the branch separating the Xenambulacraria from other Bilateria is short,  
279 it would still be predicted that certain characters uniting these taxa exist. Accordingly, the  
280 occurrence of neuropeptides in xenacoelomorphs related to echinoderm SALMFamides [25]  
281 has been reported previously based on immunohistochemical evidence [26, 27] to add to other  
282 known shared molecular characters [5, 28, 29].

283

284 One surprising result from our work is the lack of support for a monophyletic clade of  
285 deuterostomes when using site heterogenous models - the relationships between chordates,  
286 Xenambulacraria and protostomes are essentially unresolved. While the majority of our  
287 analyses recover a monophyletic group of chordates plus protostomes the support values are  
288 very low meaning there is no solid evidence to refute the traditional protostome/deuterostome  
289 dichotomy. All possible relationships between chordates, protostomes and Xenambulacraria  
290 are observed in different analyses (see extended info). This observation nevertheless implies  
291 an extremely short branch between the bilaterian common ancestor (Urbilateria) and the  
292 deuterostomes. If the deuterostomes are ultimately shown to be monophyletic then the short  
293 branch leading to the deuterostome common ancestor, Urdeuterostomia, suggests it should  
294 have much in common with Urbilateria. If the deuterostomes do prove to be paraphyletic then  
295 Urbilateria and Urdeuterostomia must be considered synonymous and this result has  
296 significant implications for our understanding of the characteristics of the common ancestor of  
297 Bilateria. Given that the internal branches separating the Xenambulacraria, Chordata and  
298 Protostomia are short, larger datasets and more refined methodologies (e.g. [30]) are required  
299 to adequately test the deuterostome monophyly.

300

301 **Acknowledgements:** This research was supported by European Research Council grant  
302 (ERC-2012-AdG 322790) to MJT, by the Labex TULIP (ANR-10-LABX-41) to HP, Deutsche  
303 Forschungsgemeinschaft (DFG Ha2103/4) and Johannes Gutenberg University Center of  
304 Computational Sciences Mainz (CSM) to TH, by the Swiss National Science Foundation

305 (150654) to CD, by the OIST internal fund to NS and by the JSPS Grant-in-Aid for Young  
306 Scientists (A) (JP26711022) to HN. Computations were performed at MPI Berlin, the  
307 University College London Computer Science cluster, the University College London Legion  
308 and Grace supercomputers, the Vital-IT Center for high-performance computing of the SIB  
309 Swiss Institute of Bioinformatics, and on the Mp2 and Ms2 supercomputers of the Université  
310 de Sherbrooke, managed by Calcul Québec and Compute Canada. The operation of this  
311 supercomputer is funded by the Canada Foundation for Innovation (CFI), the ministère de  
312 l'Économie, de la science et de l'innovation du Québec (MESI), and the Fonds de recherche  
313 du Québec - Nature et technologies (FRQ-NT). A rough estimation of the carbon footprint for  
314 this work is >7 tonnes of CO<sub>2</sub> for travel and > 260 tonnes CO<sub>2</sub> for computations (see STAR  
315 methods for details). We thank the staff of Sven Lovén Centre for Marine Infrastructure,  
316 University of Gothenburg for help collecting specimens.

317

318 **Author Contributions:** Initial concept: AJP, MJT, RRC, PM. Collecting specimens HN, AW,  
319 MC, PM, Sequencing of new genomes and transcriptomes: AJP, PM, MC, NS, THK, MJT,  
320 RRC, HK. Genome and transcriptome assembly and gene predictions: AJP, KJH, HK.  
321 Curation of sequence data set to remove bacterial and algal contigs: DD, MH. OMA orthology  
322 analyses: CD, BT, SM. '42' analyses and alignment cleaning: HP. Phylogenetic analyses: HP,  
323 MJT, AJP, CD, PHS. Additional comparative analysis of sequence data: RRC, MRE, MLR,  
324 MTC, TH, FM. Drafting manuscript: MJT, HP. Commenting on m/s: all authors.

325 **Declaration of interests.**

326 The authors declare no competing interests.

327

328

329 **Figure 1. Support for Xenambulacraria is strengthened in experiments designed to**  
330 **reduce systematic errors.** A. Full data set using all 1,173 genes and 350,088 positions  
331 shows limited support (60% of Jackknife replicates highlighted in red) for a sister group  
332 relationship between Xenacoelomorpha and Ambulacraria (Xenambulacraria). B. Summary  
333 figure of result in 1A. C. Full data set with long branched Acoelomorpha removed, results in  
334 increased support for Xenambulacraria (81% jackknife support). D. Data set of all species  
335 and the best 25% of genes (as measured by their ability to reconstruct known monophyletic  
336 groups) results in increased support for Xenambulacraria (94% jackknife support).  
337 Monophyletic deuterostome clade is not supported though the position of the Chordata is not  
338 reliably resolved in any analysis. All analyses used 50 Jackknife replicates (support values  
339 shown to right of nodes) analysed with PhyloBayes using the CATGTR+Gamma model.  
340 Major clades are indicated with corresponding colours. Jackknife proportions = 100% unless  
341 shown. The outgroups are shown in green. See also Figures S1-4 and Table S1.

342

343 **Figure 2. Best genes and best fitting model support Xenambulacraria hypothesis under**  
344 **different conditions** (green box). Worst genes and less well-fitting model support the  
345 Nephrozoa hypothesis (red box). Summary trees with jackknife support values shown for  
346 relationships between key clades for different methods of analysis. Best genes were selected  
347 by their ability to reconstruct known monophyletic groups. Top row analysed with better fitting  
348 site heterogeneous CATGTR+Gamma model. Bottom row analysed with less well-fitting site  
349 homogenous GTR+Gamma model. 'Dayhoff6' used Dayhoff recoding to reduce compositional  
350 bias. 'No Acoel' excluded long branched Acoelomorpha. 'All' included all species with no data  
351 recoding. Chords = Chordata, Proto = Protostomia, Ambula = Ambulacraria, Xenacoels =  
352 Xenacoelomorpha. JP= Jackknife Proportion. BP = Bootstrap proportion.

353

354 **Figure 3. Dayhoff recoding to reduce compositional bias and saturation increases**  
355 **support for Xenambulacraria.** PhyloBayes jackknife and bootstrap analyses of all genes and  
356 all taxa using CATGTR and Dayhoff recoding. The jackknife topology is shown though the

357 bootstrap topology was identical and branch lengths were almost identical. Jackknifing used  
358 50 replicates of 30,000 amino acids. Jackknife proportions (first number) and bootstrap  
359 proportions (second number) for nodes with less than 100% support for either measure are  
360 shown to the right of node supported. Bootstrap proportions are consistently higher,  
361 suggesting jackknifing provides a conservative measure of support. Xenambulacraria support  
362 is highlighted in red.

363

364 **Table 1. Comparisons of characteristics of best and worst quarters of genes from the**

365 **three data sets.** For the data from this study, from Cannon et al [2] and from Rouse et al [1]

366 we compare several aspects of the best and worst quarter of genes as ranked using our

367 monophyly score. The first five rows show posterior predictive tests of diversity and

368 heterogeneity of best and worst quarters of genes from the three data sets using site

369 homogenous (GTR) and heterogenous (CATGTR) models of site evolution. For all three

370 data sets and for all three tests the CATGTR model provides a closer fit to the observed

371 statistic than the site homogenous GTR model as estimated by the z-score shown here.

372 There is a slightly better fit of model to data for the best genes compared to the worst genes.

373 The three tests are computed with the readpb\_mpi programme of the PhyloBayes\_mpi

374 suite: diversity (site-specific amino acid preferences), max heterogeneity (maximal compositional

375 heterogeneity observed across the taxa), and mean heterogeneity (mean squared heterogeneity

376 across taxa). The remaining rows show comparisons of best and worst genes made using

377 the CATGTR model: Congruence score measures average monophyly score per gene and

378 % recovered clades measures percentage of clans present in the super matrix LG+F+G tree

379 recovered by single genes using the same model, in all cases the best quarter are better. #

380 positions, % missing data and number of constant positions have similar values between

381 best and worst genes. Cross validation scores show how much better the CATGTR model

382 fits the data compared to the GTR model. For all data sets and partitions, trees based on the

383 best genes are consistently longer and slightly more saturated (saturation estimated as in [7])

384 from the  $\alpha_0$  parameter, using the CATGTR patristic distances) than those based on the worst  
385 genes.

386 **STAR methods**

387

388 **CONTACT FOR REAGENT AND RESOURCE SHARING**

389 Further information and requests for resources and reagents should be directed to and will  
390 be fulfilled by the Lead Contact, Max Telford (m.telford@ucl.ac.uk).

391

392 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

393 *Xenoturbella bocki* were collected from mud dredged at approx. 60 metres depth in  
394 Gullmarsfjord, Sweden.

395 *Symsagittifera roscoffensis* were collected from intertidal regions of beaches in region of  
396 Roscoff, France.

397 *Meara stichopi* were collected by dissection from the pharynx of the sea cucumber

398 *Stichopus* sp. The sea cucumbers were collected in the sea close to Bergen, Norway.

399 *Pseudaphanostoma variabilis* were found in sediment collected close to the island of Hållö  
400 close to Smögen, West coast Sweden.

401 *Praesagittifera naikaiensis* were collected in sediment dredged from the sea bed close to  
402 Onomichi, Hiroshima, Japan

403 *Paratomella rubra* were collected from intertidal sands of Filey bay, Yorkshire, United  
404 Kingdom.

405 *Isodiametra pulchra* came from a lab strain from the University of Innsbruck, Austria.

406

407 **METHOD DETAILS**

408 ***Xenoturbella bocki* genome**

409 Independent Illumina fragment libraries were made from two single animals, which had been  
410 starved for at least 3 months in the presence of Penicillin, Streptomycin and Gentamycin  
411 antibiotics to minimize environmental bacterial contaminations. The fragment libraries had  
412 insert sizes of ~200bp and ~150 bp and were sequenced as single paired reads with read

413 length of 36-100bp. Overlapping paired reads were joined using flash [31]. The majority of  
414 sequences were obtained from these two libraries of which 10 lanes were sequenced.  
415  
416 Mate pair libraries were constructed from DNA isolated from several animals with insert  
417 sizes of 700, 1,000, 1,500 and 2,000 bp. After standard Illumina filtering all sequences  
418 shorter than 31bp were discarded. All reads were subsequently filtered for adaptor  
419 sequences, PCR duplicates and quality with SOAPfilter\_v2.0  
420 ([https://github.com/tanghaibao/jcvi-bin/blob/master/SOAP/SOAPfilter\\_v2.0](https://github.com/tanghaibao/jcvi-bin/blob/master/SOAP/SOAPfilter_v2.0)) using standard  
421 settings except setting the insert sizes and the appropriate ascii quality shifts. A total of  
422 731,057,046 reads were assembled simultaneously using SOAPdenovo (v2) [32] using  
423 settings -K 31 -M3 -F -U -g200. A total of 108,063,238 bp were assembled in a total of  
424 21,594 scaffolds. The average scaffold length was 5004 bp, the longest scaffold had a size  
425 of 317,597 bp. Including contigs not merged into scaffolds the total sequence size was  
426 119,097,168 bp with an average length of 1210 bp an N50 of 22,208 and an N90 of 443bp.  
427 Additional gaps were filled using SOAP Gapcloser v1.12  
428 (<http://soap.genomics.org.cn/soapdenovo.html>).

429

430 Using the human matrix, Genescan [33] was used to generate predictions of coding regions  
431 resulting in 23 Mb of protein coding sequence (N50: 1872 bp) in 21,769 predicted protein or  
432 peptide sequences, which were subsequently used for phylogenomic analyses.

433

#### 434 ***Symsagittifera roscoffensis* genome**

435 A standard fragment Illumina library was made from a pool of symbiont free hatchlings,  
436 which were raised in artificial sea water in the presence of antibiotics. Reads were  
437 processed as described for *Xenoturbella* above. 526,232,442 reads were assembled using  
438 SOAPdenovo2 (-M3, -R, -d1, -K31) and the Celera assembler using the settings for large  
439 and heterozygous genomes. Single gene analyses indicated that the two assemblers had  
440 different qualities in different regions of the genome. Hence the entire Soap assembly and

441 the Celera assembly using its contigs and degenerate contigs larger than 500 bp were jointly  
442 assembled using minimus2 [34]. Although the total assembled genome size of about 1 Gbp  
443 from the SOAPdenovo assembly was reduced to about 450 Mb of assembled sequence  
444 many single gene analyses and PCR amplifications indicated that many more genes are  
445 represented in the joint assembly in significantly longer gene models. The joint assembly  
446 had an N50 of 2,905bp and a N90 of 587bp. Analysis of missing sequences indicated that  
447 most of the removed part is composed of repetitive sequence. The total number of  
448 predictions for coding sequences is 113,993 and comprising a total of 52Mb. A  
449 transcriptome was also sequenced from *S. roscoffensis* mixed stage embryos using  
450 standard methods.

451

#### 452 **Amplifying genomes of small acoels**

453 Due to their small sizes one whole animal each of *Meara stichopi*, *Nemertoderma westbladi*,  
454 and *Pseudaphanostoma variabilis* were used without prior DNA extraction to directly amplify  
455 genomic DNA using the illustra Genomphi V2 DNA amplification Kit (GE Healthcare Nr.: 25-  
456 6600-30). Amplified DNA was cleaned by Isopropanol precipitation and sheared to 1.5-3 kb  
457 fragments using speed code SC6 on the Hydroshear DNA Shearing Device (Thermo Fisher  
458 Scientific). After additional cleaning and quantification 1 micrograms DNA from each animal  
459 was used to generate standard illumina fragment libraries and these were sequenced as  
460 paired end with sequence length 100 bp. Sequence data have been submitted to the  
461 European Nucleotide Archive (ENA) under accession number PRJEB25577.

462

463 *Nemertoderma westbladi* was collected from mud at the site "Telekabeln" in the  
464 Gullmarsfjord in July 2009. For *Nemertoderma westbladi*, 800,863,374 reads equalling ~80  
465 Gb of sequence were used for the genome assembly using SOAPdenovo2. The best results  
466 were obtained using the settings -K39 -d0 -M 3 -map 45. The assembly comprised about  
467 205 Mb with an N50 of about 380 bp. 80,966 gene predications resulted in 38Mb of coding  
468 sequence.

469

470 For *Meara stichopi* 1,167,743,394 reads (~110 Gb) were read. An assembly was generated  
471 using standard settings and `-K -M 3`. The assembly had a total size of about 1.4 Gbp and  
472 an N50 of 1.1 Kb. A total of 130,115 protein or peptide fragments were predicted comprising  
473 37Mb of coding sequence.

474

475 *Pseudaphanostoma variabilis* was collected from shell gravel near the island Hällö close to  
476 Smögen in July 2009. The *Pseudaphanostoma variabilis* genome was assembled from  
477 672,950,533 reads with the SOAPdenovo2 settings `-K 31 -d 0 -M 3 -map 36` and resulted  
478 in an assembly size of about 413 Mb. 115,245 gene predictions comprised 45 Mb of coding  
479 sequence.

480

481 The *Praesagittifera naikaiensis* genome was sequenced and assembled at the Okinawa  
482 Institute of Science and Technology. 1,148,317 sequences with a total size of about 1.2 Gb  
483 and an N50 of 4,452 bp resulted in 400,106 gene predictions comprising 233Mb of coding  
484 sequence.

485

#### 486 ***Paratomella rubra* transcriptome**

487 Specimens of the acoel *Paratomella rubra* were collected from intertidal sand in Filey Bay,  
488 Yorkshire, UK. RNA was prepared and sequenced, the transcriptome was assembled and  
489 cross-contaminants were removed and proteins predicted as described in (Egger et al 2015  
490 [18]). Data available in the NCBI Short Read archive: SRX3470480.

491

#### 492 ***Isodiametra pulchra* transcriptome**

493 Specimens of the acoel *Isodiametra pulchra* were harvested from a lab stock provided by B  
494 Egger, Innsbruck. RNA was prepared and sequenced, the transcriptome was assembled and  
495 cross-contaminants were removed and proteins predicted as described in ref [18]. Data  
496 available in the NCBI Short Read archive SRX3469680.

497

#### 498 **Initial contaminant cleaning**

499 All sequences were scanned for contaminating bacterial sequences using the PhymmBL  
500 program [35]. Sequences were additionally clustered based on tetranucleotide frequencies  
501 using an emergent self-organizing map (ESOM).

502

#### 503 **Removing redundancy**

504 We translated gene predictions from genomes and transcriptomes into protein sequence  
505 and, when both present from a given species, we joined both predictions and clustered using  
506 CD-HIT with a 97% identity threshold [36], resulting in non-redundant proteomes for each  
507 species. We obtained 32,456 complete gene predictions in *Symsagittifera roscoffensis*,  
508 35,867 complete gene predictions in *Meara stichopi*, 23,233 complete gene predictions in  
509 *Nemertoderma westbladi*, 27,378 complete gene predictions in *Pseudophanostoma*  
510 *variabilis*, 24,329 complete gene predictions in *Paratomella rubra*, 19,206 complete gene  
511 predictions in *Xenoturbella bocki*.

512

#### 513 **Initial ortholog predictions using OMA**

514 Non-redundant peptide datasets from 67 species including 9 Xenacoelomorpha species, 8  
515 Chordata, 15 Ambulacraria, and 13 Protostomia and 22 non-Bilateria organisms were  
516 processed by the OMA standalone software version 0.99w [37], using default settings. This  
517 identified 245,524 Orthologous Groups (OGs)—sets of genes in which all members are  
518 orthologous to all other members. From these, we selected the 3,683 OGs which had a  
519 minimum of 34 species represented (at least 50% of all species), and further filtered 1,665  
520 OGs containing at least one member of Xenoturbellida and Nemetodermatida and Acoela.

521

#### 522 **Reducing missing data, adding species and initial cleaning using 42 software**

523 Transcriptomic data from 77 species were then incorporated into the 1,665 previously  
524 assembled core orthologous clusters using a multiple Best Reciprocal Hit approach

525 implemented in the newly designed Forty-Two software  
526 (<https://bitbucket.org/dbaurain/42/downloads>). First, we removed the most divergent  
527 sequences, which are the most likely to be paralogs or contaminants. More precisely for  
528 each species having multiple sequences, each sequence was BLASTed against the rest of  
529 the alignment and the best hit identified; a sequence was removed if it overlapped with the  
530 best hit sequence by  $\geq 95\%$  and if its BLAST score was below the best hit score by a given  
531 threshold. Using a threshold of 10%, 17,480 sequences were removed. The resulting  
532 clusters were cleaned using HmmCleaner version 1.8 [38] and the same process was  
533 repeated, this time removing 4,267 additional sequences. Most of these sequences were  
534 sequencing variants of the same transcripts (due to sequencing errors or to *in vivo* transcript  
535 degradation).

536

### 537 **Removing potential contaminants**

538 As in Simion et al. [9], alignments of ribosomal proteins containing a large eukaryotic  
539 taxonomic diversity were used to detect contaminations. We used BLASTP against several  
540 custom databases to detect and remove the contaminants. An additional screening was  
541 done using BLASTN to remove the few remaining contaminants from *Homo sapiens* and  
542 *Danio rerio*. The case of homoscleromorph and calcareous sponges was analysed  
543 differently, because of the absence of clean complete genomes that can serve as a  
544 reference for decontamination. For each alignment, we BLASTed each poriferan sequence  
545 against the other sequences and removed the 2,434 sequences that had a BLAST bit score  
546 to the 'wrong' clade that was 5% higher than to the expected clade (i.e., Calcarea,  
547 Demospongiae, or Homoscleromorpha).

548

549 To discard genes for which orthology/paralogy relationships are difficult to infer, we made  
550 alignments using Mafft [39] (mafft --quiet --localpair --maxiterate 5000 --reorder), cleaned  
551 alignments with HmmCleaner and constructed RAxML trees [40] using the LG+Gamma+F  
552 model. We then computed the number of taxonomic groups (among the 14 clades displaying

553 a long basal branch: Acoela, Anthozoa, Calcarea, Chordata, Demospongiae, Ecdysozoa,  
554 Echinodermata, Hemichordata, Homoscleromorpha, Lophotrochozoa, Medusozoa,  
555 Nemertodermatida, Rotifera and Xenoturbellida) displaying paralogous copies (see [9]) and  
556 eliminated the 157 genes with  $\geq 5$  cases of paralogy.

557

558 To reduce the amount of missing data and the computational burden, we removed 21  
559 species (highly incomplete, taxonomically redundant or fast-evolving) and then the 137  
560 genes in which more than one of the following 8 groups (Acoela, Nemertodermatida,  
561 Xenoturbellida, Echinodermata, Hemichordata, Chordata, Protostomia and outgroup) is  
562 missing. We had three criteria for choosing which taxa to retain: 1. Taxonomic diversity with  
563 the aim of picking a member of each of the major groups of a given clade (i.e. not all  
564 arthropods for Ecdysozoa). 2. Avoiding taxa with known issues such as extreme branch  
565 lengths or compositional biases (e.g. picking a shorter branch nematode rather than the  
566 familiar but rapidly evolving *Caenorhabditis elegans*). 3. choosing a species with fewest  
567 missing data.

568

569 Our last quality check was based on the rationale that non-orthologous sequences (being  
570 either a contaminant or a paralog and thus misplaced) typically display very long branches  
571 when constrained on the species tree. First, alignments were cleaned with HmmCleaner  
572 version 1.8 [38] and BMGE [41], and concatenated using SCAFoS [42]. The phylogeny  
573 inferred using RAxML [40] from the supermatrix under the LG+Gamma<sub>4</sub>+F model was  
574 considered as a proxy of the species tree (note that xenacoelomorphs were sister to all other  
575 bilaterians in this tree). Then, for each alignment, the reference topology was pruned of the  
576 species missing in that alignment, and branch lengths on this constrained topology were  
577 estimated using RAxML (LG+Gamma<sub>4</sub>+F model). This allowed us to compare terminal  
578 branch lengths observed in the single-gene tree to those observed in the pruned supermatrix  
579 tree, and to remove sequences for which the branch-length ratio was  $> 5$ , hence eliminating  
580 642 questionable sequences.

581 Finally, we only kept the 1173 alignments in which at most 16 species were missing. We  
582 used SCaFoS to assemble the supermatrix, build chimeras of closely-related species  
583 (*Oscarella carmelae*/*Oscarella* SN2011, *Saccoglossus kowalevskii*/*Saccoglossus*  
584 *mereschkowskii* and *Cephalodiscus gracilis*/*Cephalodiscus hodgsoni*) and retained only the  
585 slowest-evolving sequence when multiple copies were available for a given species (using  
586 Tree-Puzzle and the WAG+F model to compute distances). This produced a supermatrix  
587 containing 350,088 amino acid positions for 59 species, with an overall amount of 23.5%  
588 missing data.

### 589 **Dataset quality**

590 To compare of our dataset with those of Cannon and Rouse [1, 2], for each gene separately  
591 we computed a phylogeny using RAxML (LG+Gamma<sub>4</sub>+F model) [40]. We then computed  
592 the number of tree bipartitions observed in the supermatrix tree (constructed with the same  
593 model) that are recovered by each gene. We assume that the majority of partitions in the  
594 supermatrix tree are likely to be correct and the percent of recovered bipartitions in the  
595 single gene trees is thus an estimation of dataset quality. Dataset quantity was measured as  
596 total amino acids.

597

### 598 **Phylogenetic inference**

599 The supermatrix was analysed with the site-heterogeneous CATGTR model [43] using  
600 PhyloBayes-MPI version 1.8 [44] after the removal of constant positions ('-dc' option) and  
601 with the site-homogeneous GTR model using raxml version 8.2.8 [40]. The use of LG or  
602 LG4X models gave virtually the same results as GTR. The robustness of phylogeny was  
603 inferred with 100 rapid bootstraps in the case of the GTR model and with 100 gene  
604 jackknives in the case of the CATGTR model.

605

### 606 **Stratifying genes according to support for known monophyletic groups**

607 To select the genes from all three data sets (this study, Rouse et al. and the larger 881  
608 genes data set of Cannon et al.) most likely to contain easy to extract phylogenetic signal,  
609 we used two different approaches. First, each gene was analysed separately to find their  
610 individual level of support for known monophyletic groups. All Xenacoelomorph sequences  
611 were removed such that the monophyly measure was independent of the presence of this  
612 clade. For each aligned and trimmed gene, a tree was reconstructed using phymI [45]  
613 (settings -d aa -o tlr -a e -c 5). Each resulting tree was analysed using a custom perl script  
614 that measured the support for the following uncontroversial monophyletic groups: Cnidaria,  
615 Ambulacraria, Hemichordata, Echinodermata, Chordata, Ecdysozoa, Lophotrochozoa,  
616 Porifera, Ctenophora (where present) Protostomia and Bilateria The monophyly score for  
617 each clade was calculated as the size of the largest clade on the tree containing species  
618 from the monophyletic group in question divided by the total number of species from that  
619 monophyletic group in the dataset. For example, if there were five chordates in the data set  
620 and the largest chordate-only grouping on the tree contained four of them, the monophyly  
621 score for chordates would be  $\frac{4}{5} = 0.8$ . The total score for the tree was calculated as the  
622 monophyly score averaged over all clades. Clades with fewer than two species in the tree  
623 were ignored. The data sets were then ranked by monophyly score and concatenated (with  
624 Xenacoelomorphs now included) in order from best (highest monophyly score) to worst.

625

626 For each of the three stratified data sets (ours, Cannon et al. and Rouse et al.) we took the  
627 genes representing the first 25% of positions (best) and the last 25% positions (worst)  
628 and performed jackknife resampling to produce 50 jackknife replicates each containing  
629 ~30,000 positions. Each jackknife replicate data set was analysed using PhyloBayes-MPI  
630 and a CATGTR+Gamma model with a single run and stopping after 1500 cycles. The  
631 jackknife summary tree was produced using a bpcomp analysis using all 50 replicates with a  
632 burnin discarding the first 1000 cycles. We also inferred Maximum Likelihood trees using the  
633 GTR+Gamma model with RAxML [40] based on the concatenations of the best and worst  
634 25% of genes.

635

636 In a second closely related approach, we sorted the genes according to the percentage of  
637 bipartitions observed in the supermatrix tree that are recovered by each gene and took the  
638 25% genes with the highest (lowest) values as the best (worst) genes this time including all  
639 species. These approaches gave congruent results and we present only those from the first  
640 approach.

641

#### 642 **Dayhoff recoding**

643 This was performed using the “-recode Dayhoff6” command in PhyloBayes-MPI.

644

#### 645 **Posterior Predictive Analyses (ppred)**

646 These were conducted using PhyloBayes ppred command as described in ref [19].

647

#### 648 **Carbon footprint calculations**

649 The carbon footprint for travel was computed only for flights for the three meetings  
650 specifically organised for this project, so constitute a small underestimate. We used the  
651 calculator of the International Civil Aviation Organization ([https://www.icao.int/environmental-  
652 protection/Carbonoffset/Pages/default.aspx](https://www.icao.int/environmental-protection/Carbonoffset/Pages/default.aspx)), which did not include radiative forcing, so  
653 seriously underestimating the impact on global warming (Table S2).

654 The carbon footprint for computation was more difficult to compute since analyses were  
655 done in multiple labs, using various computers. More importantly, we did not archive all  
656 computations done for this work (e.g. preliminary analyses). We used the reasonable  
657 hypothesis that the jackknife analyses with the CATGTR model are by far the largest  
658 contributor and compute their footprint only. This certainly leads to an underestimation  
659 (ignoring for example assembly of genomes/transcriptomes, dataset building, dataset  
660 curation, RAxML analyses and Dayhoff analyses were ignored). For simplicity we also

661 assumed that all the computations were done on a single computer, mp2 of  
662 ComputeCanada (<https://wiki.calculquebec.ca/w/Accueil>).

663 For 3 taxon sampling experiments, the 100 jackknife replicates of ~90,000 positions were  
664 performed on 6 nodes of 24 cores. The average CPU time for a single replicate was 520.5  
665 hours, giving a total of 936,900 hours ( $=520.5*6*100*3$ ). The 50 jackknife replicates of  
666 ~30,000 positions were performed on 2 nodes of 24 cores, for 3 datasets (Our data, Cannon  
667 and Rouse), 2 taxon samples, 2 data samples (best/worse) and 2 methods. The average  
668 time for a single replicate is 188.8 hours, so a total of 453,120 hours of a single node  
669 ( $=188.8*2*50*3*2*2*2$ ). Total time for all jackknife experiments assuming a single node is  
670 1,390,020 hours.

671 A node of mp2 consumes 300 W, to which we add cooling (22,75%) and other components  
672 (~5%) (Suzanne Talon, personal communication), so one hour of computation corresponds  
673 to ~0.38 kWh ( $=0.3*1.2775$ ). Total electric energy consumption for our CATGTR jackknife  
674 replicates was 531,683 kWh ( $=1,390,020*0.38$ ). To convert this into CO<sub>2</sub> emissions, we used  
675 the world average carbon intensity of power generation in 2017  
676 (<https://www.iea.org/tcep/power/>), 491 gCO<sub>2</sub>/kWh, which leads to an estimate of 261 tonnes  
677 of CO<sub>2</sub> ( $=531,683*0.000491$ ).

678

## 679 **QUANTIFICATION AND STATISTICAL ANALYSIS**

### 680 **Jackknife procedure and tests for reliability.**

681 A jackknife replicate was generated by randomly sampling single-gene alignments without  
682 replacement until >90,000 positions (~390 genes per replicate for most) or >30,000 positions  
683 (~130 genes per replicate for the analyses of best and worst genes) depending on analysis  
684 were selected. For PhyloBayes-MPI analysis of jackknife replicates, 3000 cycles were  
685 performed and consensus tree and jackknife support were obtained as in Simion et al.  
686 (2017).

687

688 To see whether the number of cycles gives an accurate measure, we experimented by  
689 extended our chains. Increasing the number of cycles did not alter jackknife proportions  
690 (Table S1.).

691

692 Similarly, running two chains of each jackknife replicate until convergence also strengthens  
693 our results. We performed an experiment where we ran two chains for each of 100 jackknife  
694 samples of 30k positions for the 'best' quarter of positions of our data with all taxa. Of these,  
695 51 pairs of chains converged (maxdiff <0.3) and 49 pairs did not (maxdiff > 0.3) - we  
696 compared the results from converged and imperfectly converged sets (Table S1.).

697

698 50 of 59 nodes received 100% support (Jackknife Proportion JP = 100%) in both converged  
699 and non-converged datasets and all but 4 received >90% support in both converged and  
700 non-converged pairs of chains. For all nodes that did not receive maximum support, the level  
701 of support is very similar for the converged and the imperfectly converged set. Interestingly,  
702 for 7 out of 9 nodes, the level of support in the converged set of runs was higher.  
703 Xenambulacraria support increased from 0.91 to 0.96. Chordata + Protostomia from 0.45 to  
704 0.58. Only support for monophyly of Acoelomorpha and sister-group of *Ircinia* and *Chondrilla*  
705 was lower in the converged data (0.5 and 0.98) than in non-converged (0.65 and 1).

706

707 We also compared the results from Jackknifing to those from Bootstrapping (which uses full  
708 sized data sets as opposed to jackknifing which uses a smaller subsample). Bootstrapping  
709 can be applied in some of the less CPU intensive analyses (reduced alphabet analyses  
710 which are significantly quicker). When we do this (100 replicates) for our full data set with all  
711 species, the supports were very similar to those of the jackknife based on 90K positions,  
712 and, as expected, slightly higher (see below). Interestingly, the support value for  
713 monophyletic Xenambulacraria increases from 90% jackknife to 98% bootstrap support  
714 (Table S1.). This supports our contention that jackknifing provides a conservative estimate of  
715 support.

716

717 Due to the relatively small size of the main Cannon et al. [2] data set (~45k positions) we  
718 managed to run a full PhyloBayes analysis to convergence on a complete data set. We used  
719 the CATGTR site heterogeneous model on a data set from which the long branched  
720 Acoelomorpha had been removed. We found *Xenoturbella* + Ambulacraria supported with a  
721 value of 1.0 posterior probability showing that our jackknife analysis of the same was  
722 conservative (Figure S3B).

723

#### 724 **Model fit**

725 To assess the fit of different models, we performed 10-fold model cross-validations. Model fit  
726 tests were done using training data sets of 10,000 amino acids and test data sets of 2,000  
727 amino acids we used PhyloBayes version 4.1 [12] to perform cross-validation for the  
728 following models: LG+ $\Gamma$ , GTR+ $\Gamma$ , CAT+ $\Gamma$  and CAT-GTR+ $\Gamma$ . PhyloBayes was run for 1100  
729 (LG and GTR) or 3100 (CAT and CATGTR) cycles and we kept the last 1000 cycles for  
730 following likelihood computations. Cross validation was run for full data sets as well as for  
731 the best and worst genes from the gene stratification experiments. The model cross-  
732 validations in all cases clearly favoured CAT-GTR+ $\Gamma$  > CAT+ $\Gamma$  > GTR+ $\Gamma$  > LG+ $\Gamma$  (for our  
733 principal, complete data set likelihood scores with respect to LG are  $3034 \pm 152$ ,  $2270 \pm 151$   
734 and  $268 \pm 40$ ).

735

#### 736 **DATA AND SOFTWARE AVAILABILITY**

737 The sequence alignments, phylogenetic trees that support the findings of this study, as well  
738 as the script for measuring monophyletic groups, are available on GitHub  
739 (<https://github.com/MaxTelford/Xenacoelomorpha2019>). Genome and transcriptome  
740 assemblies are available at <https://figshare.com/search> project number PRJNA517079. Raw  
741 data for novel sequences are available at the Sequence Read Archive BioProject  
742 PRJNA517079.

743 **Table S1. Experiments to show that a Jackknifing approach gives conservative**  
744 **estimates of clade support. Related to Figure 1.** A. Adding more cycles makes minor  
745 differences to clade support suggesting our estimates are accurate. B. Running two chains for  
746 each replicate to convergence makes minor difference (generally slightly strengthening  
747 support for less well supported clades) suggesting our clade support estimates are  
748 conservative. B. Comparison of bootstrapping and jackknifing shows the latter is gives more  
749 conservative estimates of clade support than bootstrapping. All clades not shown in the table  
750 have a support value of JP/BP = 1.

751

752 **Table S2. Calculations of CO<sub>2</sub> produced by authors travelling to meetings related to this**  
753 **work. Related to STAR methods.** For each of three meetings the origins, destinations and  
754 number of flights are shown with the approximate CO<sub>2</sub> produced in tonnes.

755

## 756 **References**

- 757 1. Rouse, G.W., Wilson, N.G., Caravajal, J., I, and Vrijenhoek, R.C. (2016). New deep-sea  
758 species of *Xenoturbella* and the position of Xenacoelomorpha. *Nature* 530, 94-97.
- 759 2. Cannon, J.T., Vellutini, B.C., Smith III, J., Ronsquist, F., Jondelius, U., and Hejnol, A.  
760 (2016). Xenacoelomorpha is the sister group to Nephrozoa. *Nature* 530, 89-93.
- 761 3. Bourlat, S.J., Nielsen, C., Lockyer, A.E., Littlewood, D.T., and Telford, M.J. (2003).  
762 *Xenoturbella* is a deuterostome that eats molluscs. *Nature* 424, 925-928.
- 763 4. Bourlat, S.J., Juliusdottir, T., Lowe, C.J., Freeman, R., Aronowicz, J., Kirschner, M.,  
764 Lander, E.S., Thorndyke, M., Nakano, H., Kohn, A.B., et al. (2006). Deuterostome  
765 phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida.  
766 *Nature* 444, 85-88.
- 767 5. Philippe, H., Brinkmann, H., Copley, R.R., Moroz, L.L., Nakano, H., Poustka, A.J.,  
768 Wallberg, A., Peterson, K.J., and Telford, M.J. (2011). Acoelomorph flatworms are  
769 deuterostomes related to *Xenoturbella*. *Nature* 470, 255-258.
- 770 6. Aguinaldo, A.M., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R.A., and  
771 Lake, J.A. (1997). Evidence for a clade of nematodes, arthropods and other moulting  
772 animals. *Nature* 387, 489-493.
- 773 7. Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Worheide, G.,  
774 and Baurain, D. (2011). Resolving difficult phylogenetic questions: why more  
775 sequences are not enough. *PLoS Biology* 9, e1000602.
- 776 8. Laurin-Lemay, S., Brinkmann, H., and Philippe, H. (2012). Origin of land plants revisited  
777 in the light of sequence contamination and missing data. *Current Biology* 22, R593-  
778 R594.

- 779 9. Simion, P., Philippe, H., Baurain, D., Jager, M., Richter, D.J., Di Franco, A., Roure, B.,  
780 Satoh, N., Queinsec, E., Ereskovsky, A., et al. (2017). A large and consistent  
781 phylogenomic dataset supports sponges as the sister group to all other animals. *Curr*  
782 *Biol* 27, 958-967.
- 783 10. Altenhoff, A.M., Glover, N.M., Train, C.M., Kaleb, K., Warwick Vesztrocy, A., Dylus, D.,  
784 de Farias, T.M., Zile, K., Stevenson, C., Long, J., et al. (2018). The OMA orthology  
785 database in 2018: retrieving evolutionary relationships among all domains of life  
786 through richer web and programmatic interfaces. *Nucleic Acids Res* 46, D477-D485.
- 787 11. Train, C.M., Glover, N.M., Gonnet, G.H., Altenhoff, A.M., and Dessimoz, C. (2017).  
788 Orthologous Matrix (OMA) algorithm 2.0: more robust to asymmetric evolutionary  
789 rates and more scalable hierarchical orthologous group inference. *Bioinformatics* 33,  
790 i75-i82.
- 791 12. Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a Bayesian software  
792 package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25,  
793 2286-2288.
- 794 13. Lartillot, N., and Philippe, H. (2008). Improvement of molecular phylogenetic  
795 inference and the phylogeny of Bilateria. *Philos Trans R Soc Lond, B, Biol Sci* 363, 1463-  
796 1472.
- 797 14. Marlétaz, F., Peijnenburg, K.T.C.A., Goto, T., Satoh, N., and Rokhsar, D.S. (2019). A new  
798 spiralian phylogeny places the enigmatic arrow worms among gnathiferans. *Current*  
799 *Biology* 29, 1-7.
- 800 15. Gavilán, B., Perea-Atienza, E., and Martínez, P. (2016). Xenacoelomorpha: a case of  
801 independent nervous system centralization?. *Philos Trans R Soc Lond B Biol Sci.* 371,  
802 1685.
- 803 16. Perea-Atienza, E., Gavilán, B., Chiodin, M., Abril, J.F., Hoff, K.J., Poustka, A.J., and  
804 Martínez, P. (2015). The nervous system of Xenacoelomorpha: a genomic perspective.  
805 *The Journal of Experimental Biology* 218, 618-628.
- 806 17. Schiffer, P.H., Robertson, H.M., and Telford, M.J. (2018). Orthonectids are highly  
807 degenerate annelid worms. *Current Biology*.
- 808 18. Egger, B., Lapraz, F., Tomiczek, B., Müller, S., Dessimoz, C., Girstmair, J., Skunca, N.,  
809 Rawlinson, K.A., Cameron, C.B., Beli, E., et al. (2015). A transcriptomic-phylogenomic  
810 analysis of the evolutionary relationships of flatworms. *Curr Biol*.
- 811 19. Feuda, R., Dohrmann, M., Pett, W., Philippe, H., Rota-Stabelli, O., Lartillot, N.,  
812 Worheide, G., and Pisani, D. (2017). Improved modeling of compositional  
813 heterogeneity supports sponges as sister to all other animals. *Curr Biol* 27, 3864-3870  
814 e3864.
- 815 20. Fritzsche, G., Bohme, M.U., Israelsson, O., Hankeln, T., and al., e. (2008). A PCR survey  
816 of *Xenoturbella bocki*: evolution of deuterostome Hox genes. *J. Exp.Zool. (Mol. Dev.*  
817 *Evol.)* 310B, 278-284.
- 818 21. Cook, C.E., Jimenez, E., and Akam, M.E. (2004). The Hox gene complement of acoel  
819 flatworms, a basal bilaterian clade. *Evol. Dev.* 6, 154-163.
- 820 22. Sempere, L.F., Martínez, P., Cole, C., Bagaña, J., and Peterson, K.J. (2007). Phylogenetic  
821 distribution of microRNAs supports the basal position of acoel flatworms and the  
822 polyphyly of Platyhelminthes. *Evol Dev* 9, 409-415.
- 823 23. Hejnol, A., and Martindale, M.Q. (2009). Coordinated spatial and temporal expression  
824 of Hox genes during embryogenesis in the acoel *Convolutriloba longifissura*. *BMC Biol*  
825 7, 65.

- 826 24. Hejnal, A., and Martindale, M.Q. (2008). Acoel development indicates the  
827 independent evolution of the bilaterian mouth and anus. *Nature* 456, 382-386.
- 828 25. Elphick, M.R. (2014). SALMFamide salmagundi: the biology of a neuropeptide family  
829 in echinoderms. *Gen Comp Endocrinol* 205, 23-35.
- 830 26. Dittmann, I.L., Zauchner, T., Nevard, L.M., Telford, M.J., and Egger, B. (2018).  
831 SALMFamide2 and serotonin immunoreactivity in the nervous system of some acoels  
832 (*Xenacoelomorpha*). *J Morphol* 279, 589-597.
- 833 27. Stach, T., Dupont, S., Israelson, O., Fauville, G., Nakano, H., Kånneby, T., and  
834 Thorndyke, M. (2005). Nerve cells of *Xenoturbella bocki* (phylum uncertain) and  
835 *Harrimania kupfferi* (Enteropneusta) are positively immunoreactive to antibodies  
836 raised against echinoderm neuropeptides. *J. MAR. Biol. Assoc. UK* 85, 1519-1524.
- 837 28. de Mendoza, A., and Ruiz-Trillo, I. (2011). The mysterious evolutionary origin for the  
838 GNE gene and the root of bilateria. *Mol Biol Evol* 28, 2987-2991.
- 839 29. Chang, Y.C., Pai, C.Y., Chen, Y.C., Ting, H.C., Martinez, P., Telford, M.J., Yu, J.K., and Su,  
840 Y.H. (2016). Regulatory circuit rewiring and functional divergence of the duplicate  
841 admp genes in dorsoventral axial patterning. *Dev Biol* 410, 108-118.
- 842 30. Dang, T., and Kishino, H. (2019). Stochastic variational inference for Bayesian  
843 phylogenetics: A case of CAT model. *Mol Biol Evol*.
- 844 31. Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to  
845 improve genome assemblies. *Bioinformatics* 27, 2957-2963.
- 846 32. Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., et al.  
847 (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo  
848 assembler. *Gigascience* 1, 18.
- 849 33. Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human  
850 genomic DNA. *J Mol Biol* 268, 78-94.
- 851 34. Sommer, D.D., Delcher, A.L., Salzberg, S.L., and Pop, M. (2007). Minimus: a fast,  
852 lightweight genome assembler. *BMC Bioinformatics* 8, 64.
- 853 35. Brady, A., and Salzberg, S.L. (2009). Phymm and PhymmBL: metagenomic  
854 phylogenetic classification with interpolated Markov models. *Nat Methods* 6, 673-  
855 676.
- 856 36. Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the  
857 next-generation sequencing data. *Bioinformatics* 28, 3150-3152.
- 858 37. Altenhoff, A., Gil, M., Gonnet, G.H., and Dessimoz, C. (2013). Inferring hierarchical  
859 orthologous groups from orthologous gene pairs. *Plos One* 8, e53786.
- 860 38. Amemiya, C.T., Alföldi, J., Lee, A.P., Fan, S., Philippe, H., Maccallum, I., Braasch, I.,  
861 Manousaki, T., Schneider, I., Rohner, N., et al. (2013). The African coelacanth genome  
862 provides insights into tetrapod evolution. *Nature* 496, 311-316.
- 863 39. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software  
864 version 7: improvements in performance and usability. *Mol Biol Evol* 30, 772-780.
- 865 40. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-  
866 analysis of large phylogenies. *Bioinformatics* 30, 1312-1313.
- 867 41. Criscuolo, A., and Gribaldo, S. (2010). BMGE (Block Mapping and Gathering with  
868 Entropy): a new software for selection of phylogenetic informative regions from  
869 multiple sequence alignments. *BMC Evol Biol* 10, 210.
- 870 42. Roure, B., Rodriguez-Ezpeleta, N., and Philippe, H. (2007). SCAFoS: a tool for Selection,  
871 Concatenation and Fusion of Sequences for phylogenomics. *BMC Evol Biol* 7(SUPPL 1),  
872 S2.

- 873 43. Lartillot, N., and Philippe, H. (2004). A Bayesian mixture model for across-site  
874 heterogeneities in the amino-acid replacement process. *Mol Biol Evol* 21, 1095-1109.
- 875 44. Lartillot, N., Rodrigue, N., Stubbs, D., and Richer, J. (2013). PhyloBayes MPI:  
876 phylogenetic reconstruction with infinite mixtures of profiles in a parallel  
877 environment. *Syst Biol* 62, 611-615.
- 878 45. Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O.  
879 (2010). New algorithms and methods to estimate maximum-likelihood phylogenies:  
880 assessing the performance of PhyML 3.0. *Syst Biol* 59, 307-321.
- 881
- 882

874 conservative. B. Comparison of bootstrapping and jackknifing shows the latter is gives more  
875 conservative estimates of clade support than bootstrapping. All clades not shown in the table  
876 have a support value of JP/BP = 1.

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Bacterial and Virus Strains		
Biological Samples		
<i>Xenoturbella bocki</i>	Gullmarsfjord, Sweden	NCBI:txid242395
<i>Symsagittifera roscoffensis</i>	Beaches off Roscoff, France	NCBI:txid84072
<i>Meara stichopi</i>	From pharynx of <i>Stichopus</i> , off Bergen Norway	NCBI:txid84115
<i>Nemertoderma westbladi</i>	Gullmarsfjord, West coast Sweden	NCBI:txid172109
<i>Pseudaphanostoma variabilis</i>	Hållö close to Smögen, West coast Sweden	NCBI:txid2510493
<i>Praesagittifera naikaiensis</i>	Onomichi, Hiroshima, Japan	N/A
<i>Paratomella rubra</i>	Sand from Filey bay, Yorkshire, UK	NCBI:txid90914
<i>Isodiametra pulchra</i>	Lab strain from Innsbruck, Austria	NCBI:txid504439
Chemicals, Peptides, and Recombinant Proteins		
Critical Commercial Assays		
Deposited Data		
Alignments, software and trees	GitHub	<a href="https://github.com/MaXtelford/Xenacoelomorpha2019">https://github.com/MaXtelford/Xenacoelomorpha2019</a>
Genome and transcriptome assemblies	<a href="https://figshare.com/search/project/number">https://figshare.com/search/project/number</a>	PRJNA517079
Raw data for novel sequences.	Sequence Read Archive BioProject	PRJNA517079
Experimental Models: Cell Lines		
Experimental Models: Organisms/Strains		
Oligonucleotides		
Recombinant DNA		
Software and Algorithms		
PhyloBayes	[44]	<a href="http://www.atgc-montpellier.fr/phylobayes">www.atgc-montpellier.fr/phylobayes</a>
Flash	[31]	<a href="http://ccb.jhu.edu/software/FLASH/index.shtml">http://ccb.jhu.edu/software/FLASH/index.shtml</a>
SOAPfilter_v2.0	[32]	<a href="https://github.com/tanghaibao/jvarkit/blob/master/SOAP/SOAPfilter_v2.0">https://github.com/tanghaibao/jvarkit/blob/master/SOAP/SOAPfilter_v2.0</a>
SOAPGapcloser v1.12	[32]	<a href="http://soap.genomics.org.cn/soapdenovo.html">http://soap.genomics.org.cn/soapdenovo.html</a>

Genescan	[33]	<a href="http://genes.mit.edu/GENSCAN.html">http://genes.mit.edu/GENSCAN.html</a>
Soapdenovo2	[32]	<a href="https://github.com/aquaskyline/SOAPdenovo2">https://github.com/aquaskyline/SOAPdenovo2</a>
minimus2	[34]	<a href="https://github.com/sanger-pathogens/circlator/wiki/Minimus2-circularization-pipeline">https://github.com/sanger-pathogens/circlator/wiki/Minimus2-circularization-pipeline</a>
PhymmBL	[35]	<a href="https://ccb.jhu.edu/software/phymmbl/index.shtml">https://ccb.jhu.edu/software/phymmbl/index.shtml</a>
CD-Hit	[36]	<a href="http://weizhongli-lab.org/cd-hit/">http://weizhongli-lab.org/cd-hit/</a>
42		<a href="https://bitbucket.org/dbaurain/42/downloads">https://bitbucket.org/dbaurain/42/downloads</a>
HmmCleaner version 1.8	[38]	<a href="https://metacpan.org/pod/HmmCleaner.pl">https://metacpan.org/pod/HmmCleaner.pl</a>
Mafft	[39]	<a href="https://mafft.cbrc.jp/alignment/software/">https://mafft.cbrc.jp/alignment/software/</a>
OMA	[37]	<a href="https://omabrowser.org">https://omabrowser.org</a>
RAxML	[40]	<a href="https://cme.hits.org/exelixis/software.html">https://cme.hits.org/exelixis/software.html</a>
BMGE	[41]	<a href="ftp://ftp.pasteur.fr/pub/gensoft/projects/BMGE/">ftp://ftp.pasteur.fr/pub/gensoft/projects/BMGE/</a>
SCaFoS	[42]	<a href="http://megasun.bch.umontreal.ca/Software/scafos/scafos.html">http://megasun.bch.umontreal.ca/Software/scafos/scafos.html</a>
Other		

Figure 1

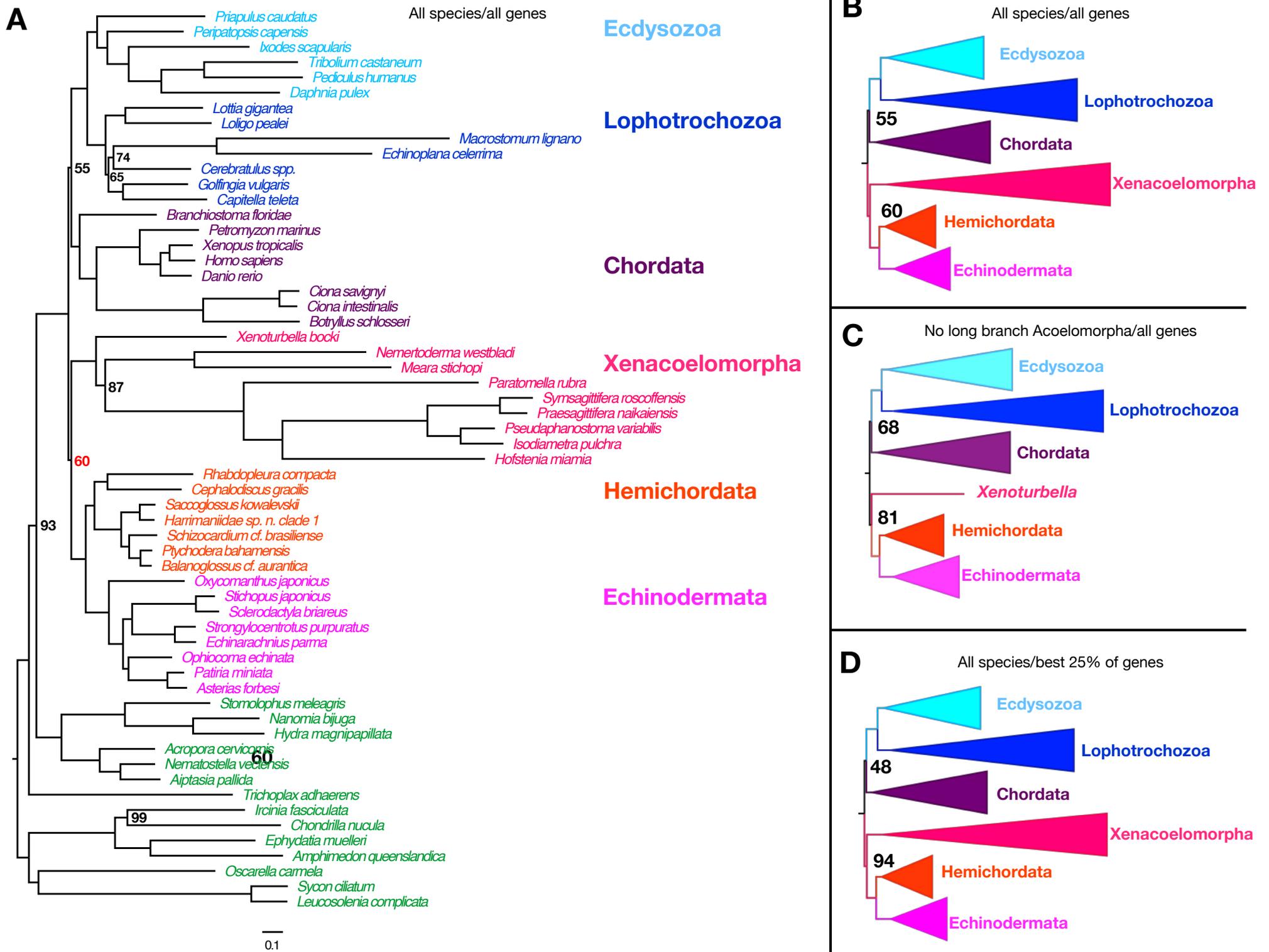
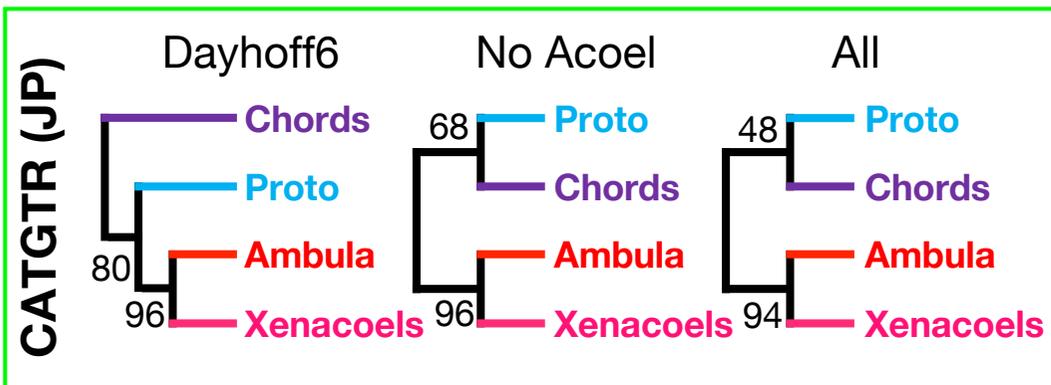


Figure 2

### BEST GENES



### WORST GENES

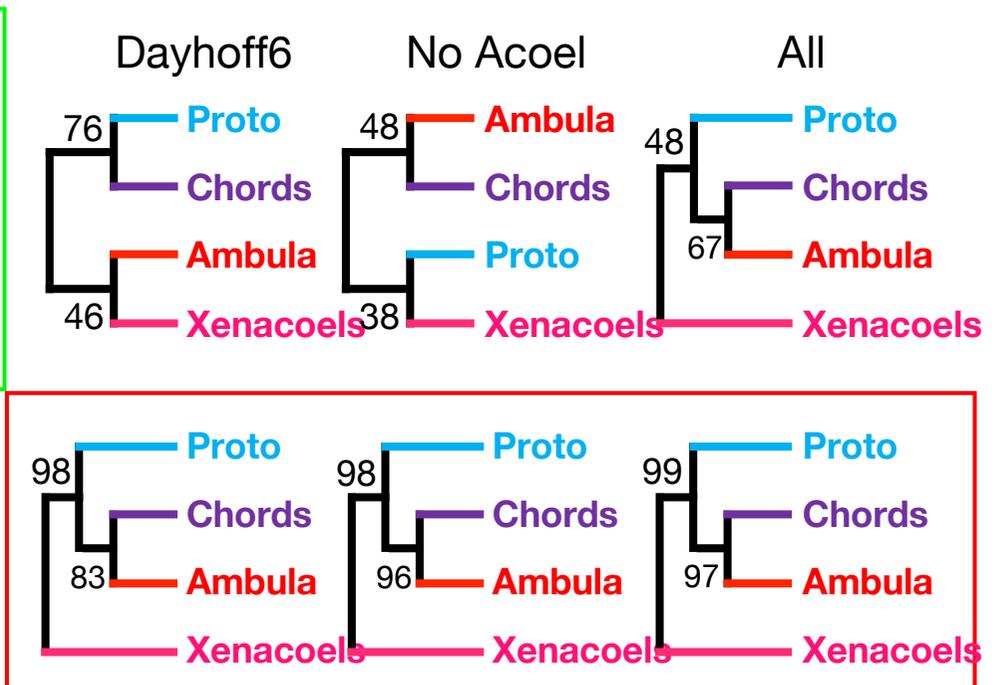


Figure 3

All species/all genes/no recoding/bootstrap

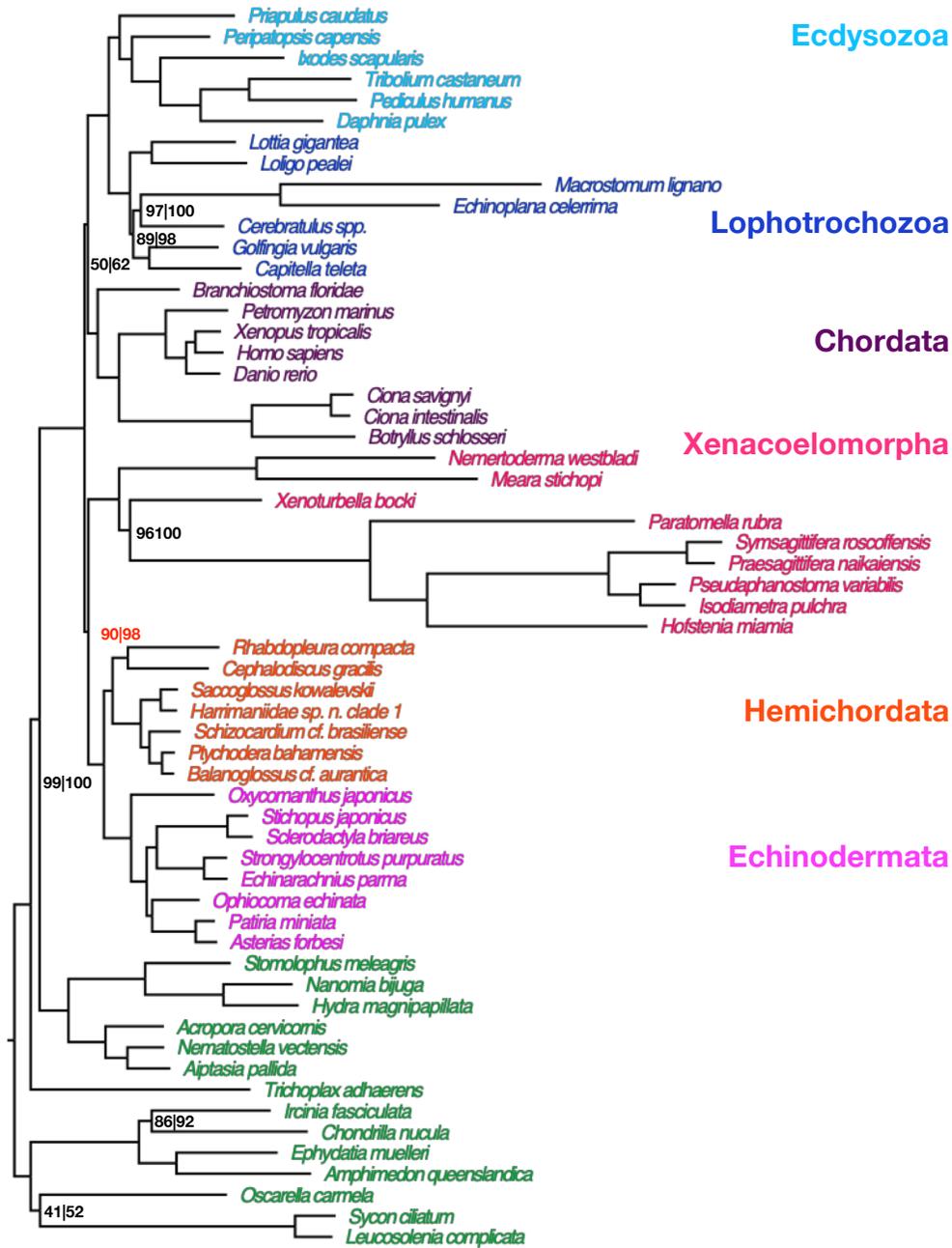


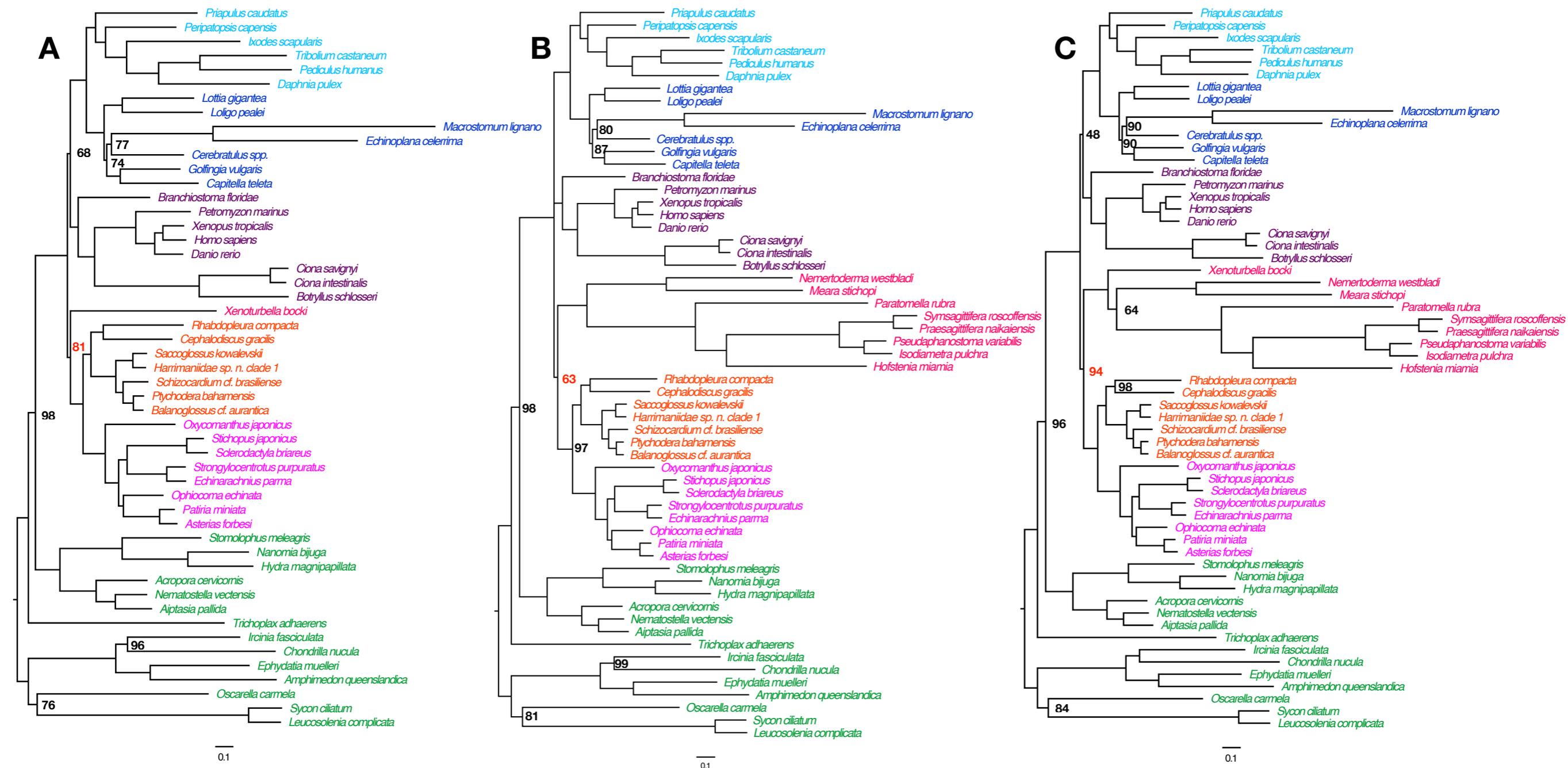
Table 1

	This study				Best
	Best Genes		Worst Genes		
model	CATGTR	GTR	CATGTR	GTR	CATGTR
diversity (Zscore)	5.7	122.0	7.0	139.7	7.8
max heterogeneity (Zscore)	17.1	37.2	88.5	197.3	9.3
mean heterogeneity (Zscore)	120.0	152.7	208.2	325.7	50.5
topology supported	X+A (94%)	X+PCA (100%)	X+PCA (48%)	X+PCA (99%)	X+A (42%)
	Best Genes		Worst Genes		Best
Congruence score	0.87		0.53		
%recovered clades	72.58		37.38		6
#positions	87791		87562		8
%missing data	24.75		22.74		3
%constant positions	20.44		24.35		1
Cross validation	2078 ± 82		3539 ± 147		291
Tree length	28.2		35.1		
Saturation	0.23		0.19		

Cannon et al. (2016)			Rouse et al. (2016)			
Best Genes	Worst Genes		Best Genes		Worst Genes	
GTR	CATGTR	GTR	CATGTR	GTR	CATGTR	GTR
132.4	10.3	199.5	3.1	69.7	3.7	84.7
12.1	43.4	106.6	1.4	1.8	2.5	4.7
68.6	169.0	276.8	6.9	7.9	29.6	39.4
X+PCA (100%)	X+P (76%)	X+PCA (100%)	X+A (50%)	X+PCA (93%)	X+PCA (50%)	X+PCA (87%)

Best Genes	Worst Genes		Best Genes		Worst Genes	
0.80	0.44		0.8		0.44	
50.45	25.17		47.40		3.47	
34276	84462		98630		98579	
39.89	36.39		43.86		40.80	
14.66	14.04		20.75		24.05	
14 ± 113	4960 ± 175		701 ± 62		997 ± 54	
50.4	63.1		27.9		31.4	
0.21	0.19		0.17		0.14	



**Supplementary Figure 1 related to Figure 1.**

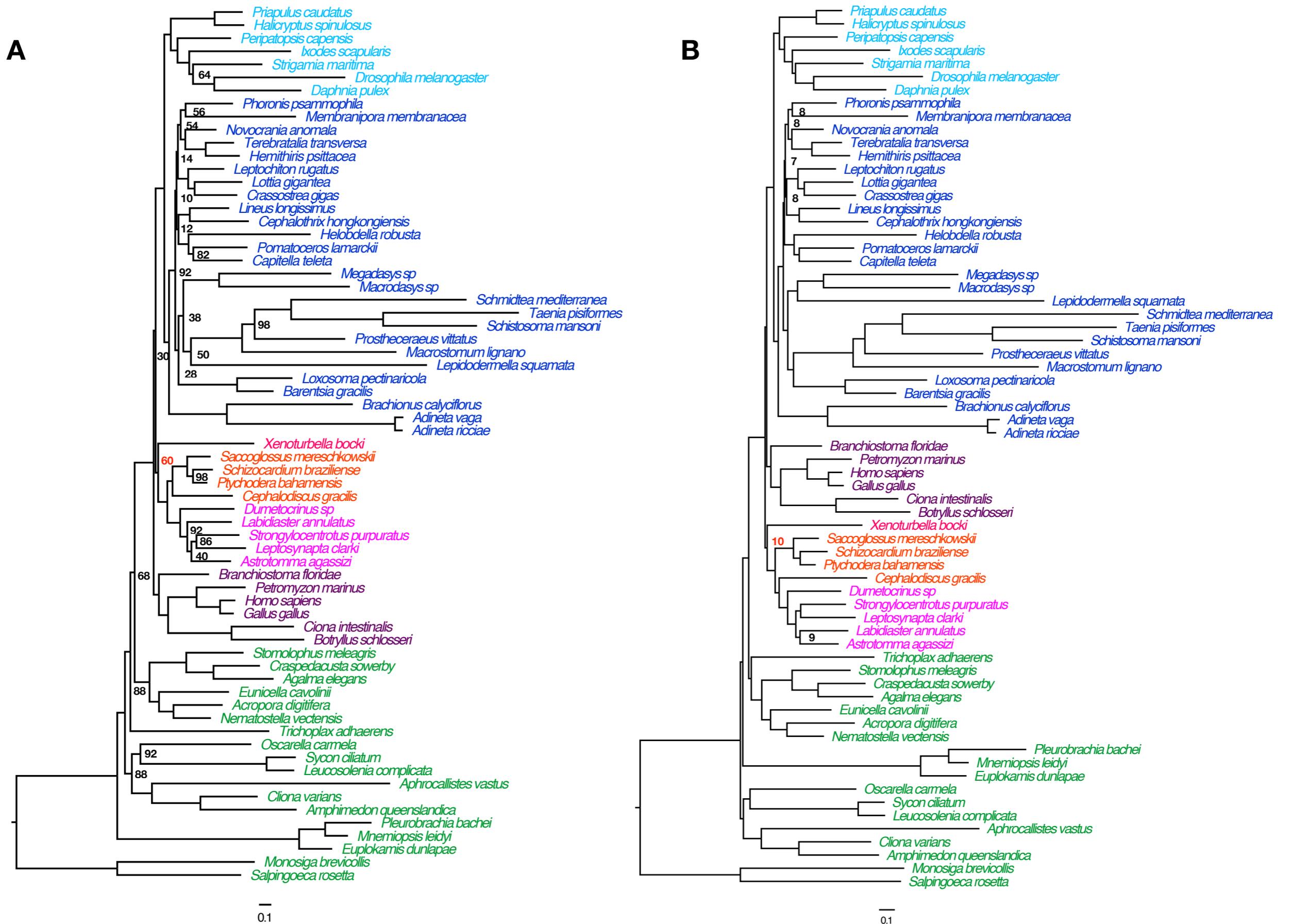
A. Phylobayes jackknife analysis. This study, All genes, No Acoelomorphs, CATGTR, 50 x 30,000 amino acids Jackknife. Jackknife proportions less than 100% shown to right of node supported.

Xenambulacraria support highlighted in red.

B. Phylobayes jackknife analysis. This study, Best quarter of genes, No *Xenoturbella*, CATGTR, 50 x 30,000 amino acids Jackknife. Jackknife proportions less than 100% shown to right of node supported.

Xenambulacraria support highlighted in red.

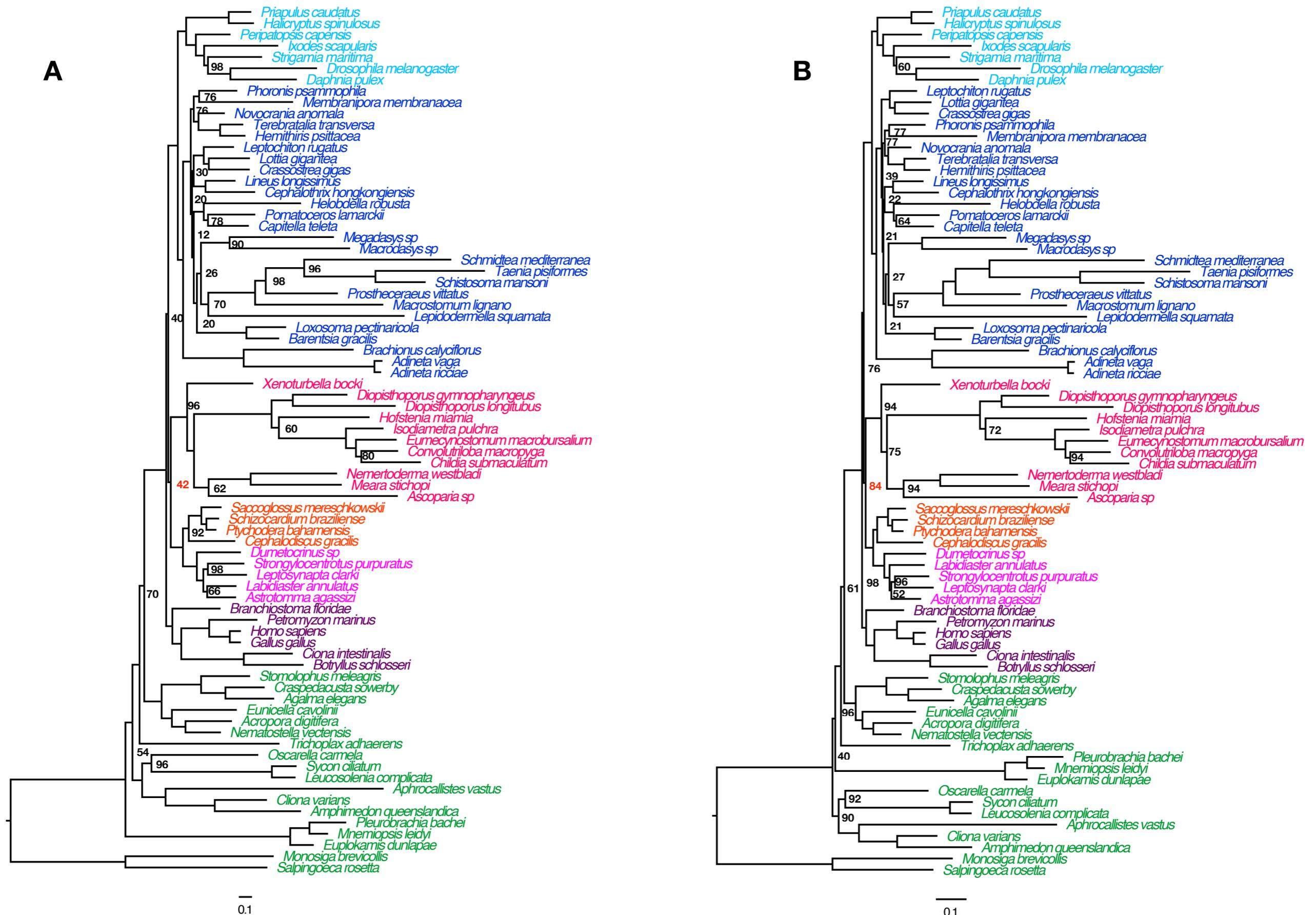
C. Phylobayes jackknife analysis. This study, Best genes, All taxa, CATGTR, 50 x 30,000 amino acids. Jackknife proportions less than 100% shown to right of node supported. Xenambulacraria support highlighted in red.



**Supplementary Figure 2 related to Figure 1.**

A. Phylobayes jackknife analysis. Cannon et al. data, All genes, No Acoelomorphs, CATGTR, 50 x 30,000 amino acids. Jackknife proportions less than 100% shown to right of node supported. Xenambulacraria support highlighted in red.

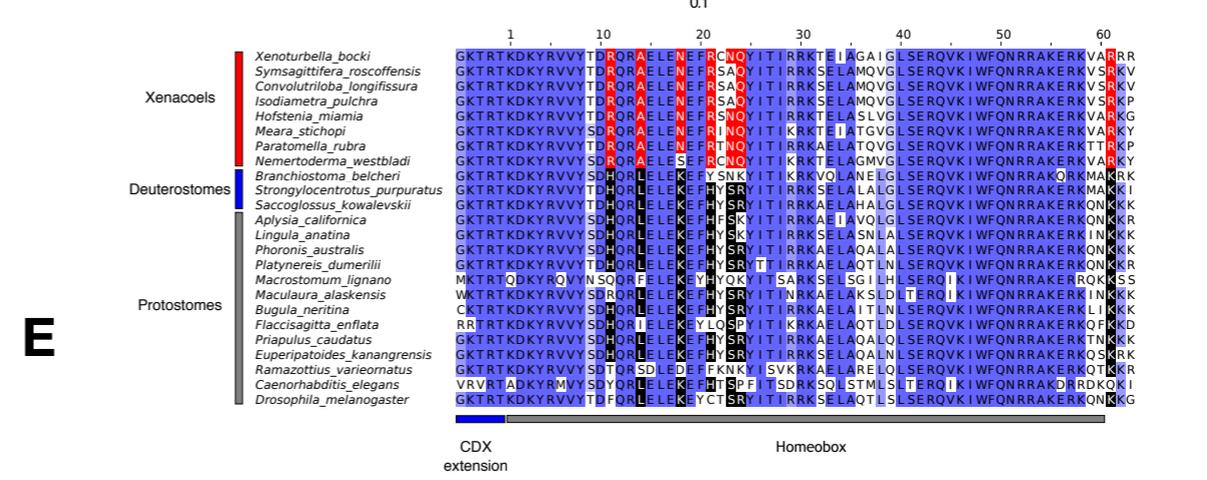
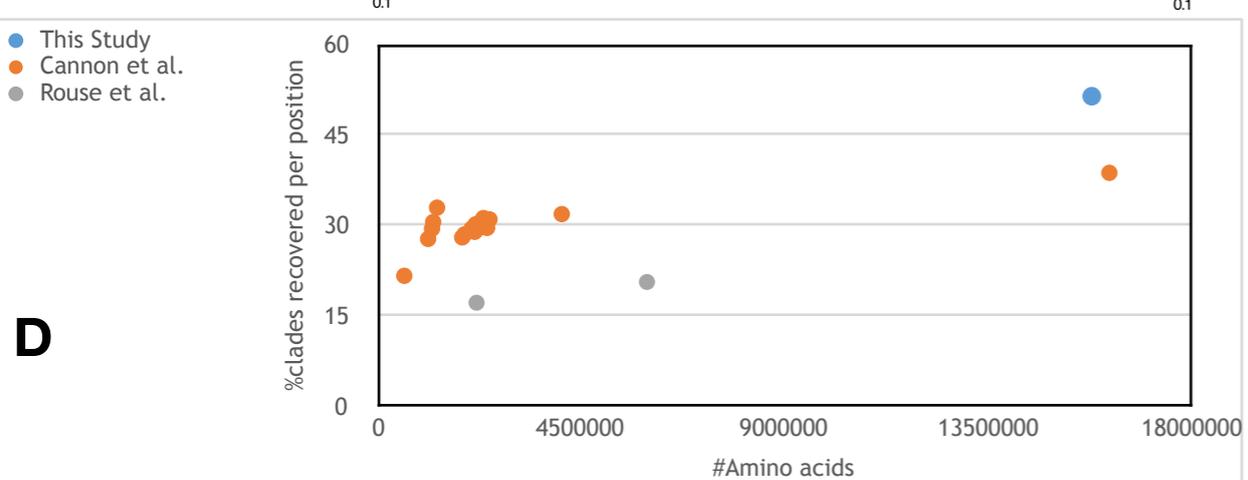
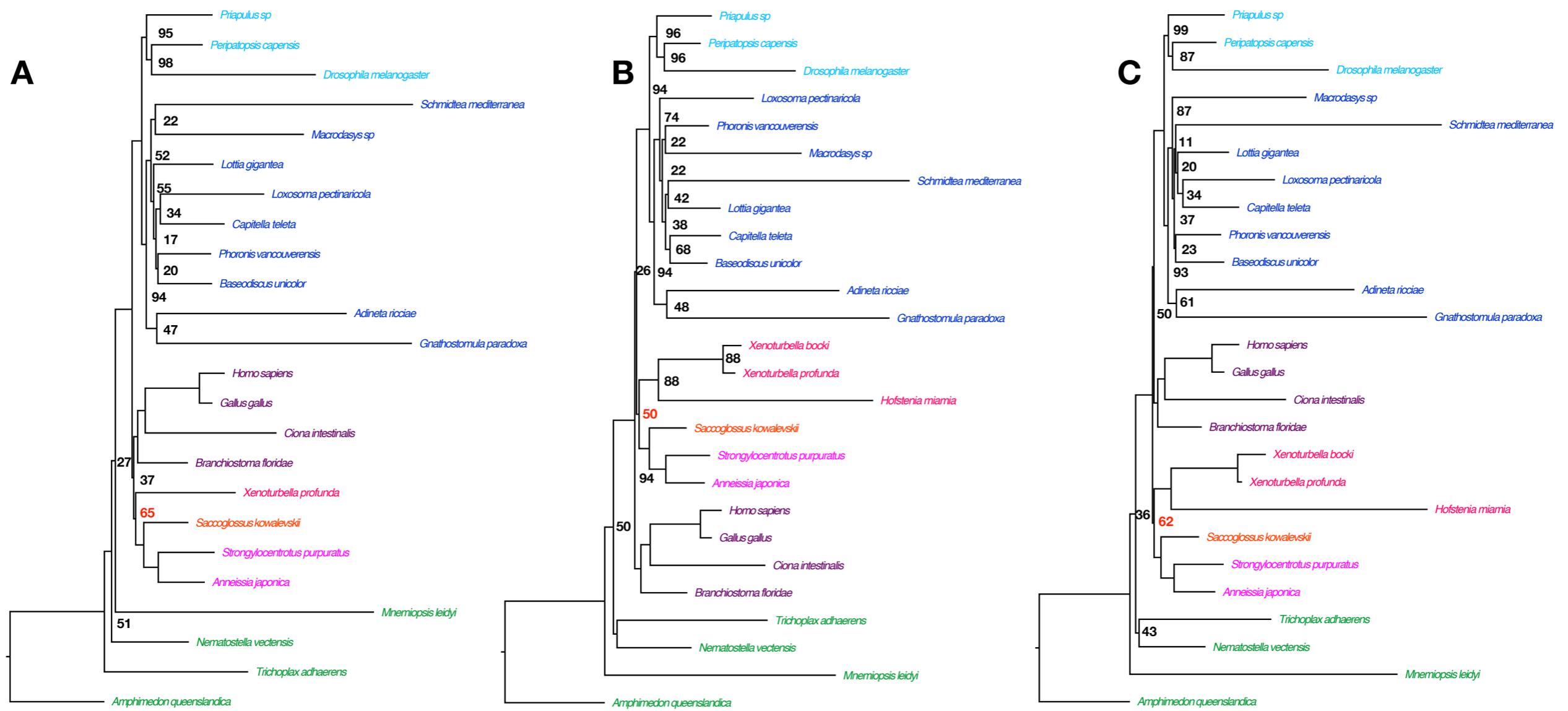
B. Phylobayes full dataset analysis. Cannon et al. data, All 212 genes, No Acoelomorphs, CATGTR. Posterior probabilities proportions less than 100% shown to right of node supported. Xenambulacraria support highlighted in red.



**Supplementary Figure 3 related to Figure 1.**

A. Phylobayes jackknife analysis. Cannon et al. data, Best genes, All taxa, CATGTR, 50 x 30,000 amino acids. Jackknife proportions less than 100% shown to right of node supported. Xenambulacraria support highlighted in red.

B. Phylobayes jackknife analysis. Cannon et al. data, All genes, All taxa, Dayhoff Recoded, CATGTR, 50 x 30,000 amino acids. Jackknife proportions less than 100% shown to right of node supported. Xenambulacraria support highlighted in red.



**Supplementary Figure 4 related to Figure 1.**

- A. Phylobayes jackknife analysis. Rouse et al. data, All genes, No Acoelomorphs, CATGTR, 50 x 30,000 amino acids. Jackknife proportions <100% shown to right of node supported. Xenambulacraria support highlighted in red.
- B. Phylobayes jackknife analysis. Cannon et al. data, Best genes, All taxa, CATGTR, 50 x 30,000 amino acids Jackknife. Jackknife proportions <100% shown to right of node supported. Xenambulacraria support highlighted in red.
- C. Phylobayes jackknife analysis. Rouse et al. data, All genes, All taxa, Dayhoff Recoded, CATGTR, 50 x 30,000 amino acids. Jackknife proportions <100% shown to right of node supported. Xenambulacraria support highlighted in red.
- D. Comparison of size and ability to reconstruct clades of different recent data sets used to reconstruct position of xenacoelomorphs. X axis: total number of amino acids in alignment. Y axis: % of clades that are present in the tree reconstructed from the total data set that are recovered by individual genes - score is the average % across genes. Cannon et al and Rouse et al presented several different data sets as shown.
- E. Alignment of homeobox region of the CDX (Caudal) gene from bilaterians. Amino acids unique to, and supporting monophyly of Xenacoelomorpha are indicated in red.