Xenacoelomorpha flatworms are basal Deuterostome

Yi Wang (✉ wyshengwuxue@126.com)

Ocean University of China at Shandong

Abstract

Background: Whether position Xenacoelomorpha as an early branch of Bilateria (Protostomes + Deuterostomes) has been intensely debated during last several decades. Considering Darwin's "tree of life", with the "Phylogenetic Species Concept", we choose mitochondrial genome as the subject to predict phylogenetic position of Xenacoelomorpha, by genes genealogy.

Results: Herein, we sequence Heterochaerus australis's mitochondrial genome and infer intrinsic relationships of Metazoan with Xenacoelomorpha. The optimal tree under the popular maximum likelihood (ML) and Bayesian phylogenetic reconstructions are consensus with each other being strongly supported. The relationship between Chordates, Ambulacrarians and Xenoturbella/Acoelomorph is resolved. To avoid previous query about alignment process, the datasets are alignmented and trimmed automatically. Reducing taxon or cutting outgroups can not affect the relationship between Xenacoelomorpha and other Metazoan. Meanwhile, analysis using CAT model and Dayhoff groups also supporting the prediction made by mtZOA, relaxing the restriction of alignment criteria (MAFFT, strategy G−ins−1, BLOSUM 62, 45, 30) introducing potential misleading signals can not challenge the tree topology indicating our auto-alignmented mitochondrial dataset is not artificially restricted one.

Conclusions: Finally, a repeatable prediction of the genes genealogy with reliable statistical support places Xenacoelomorpha as a basal Deuterostome.

Background

The "tree of life" represents the genealogy history of species and the "tree" with no doubt is metaphysically prior to any particular species concept [1]. As there is no consensus concerning the meaning of the term 'species', we choose practical one from the others, for the purpose of inferring relationships of Xenacoelomorpha with Metazoans.

There are two commonly used concept ill-named 'biological species concepts'（BSC）and 'phylogenetic species concepts' (PSC). BSC defines species as "groups of interbreeding natural populations that are reproductively isolated from other such groups." and PSC −"A species is the smallest diagnosable cluster of individual organisms within which there is a parental pattern of ancestry and descent." [2,3]. The former needs in the filed endeavor, but the later bends us to take advantage of the dramatic progressing genome data size predicting phylogenetic position of Xenacoelomorpha. Choosing PSC to predict species genealogy by using gene genealogy is the main logic of this article.

The phylogenetic species concept includes three levels: genes, organisms and species level. Same epistemology reminds us to take coalesced homolog genes genealogy history as the representative of both organisms and species genealogy once the picked genes is "majority enough" and their "tree" topology is in concordance with each other [1]. Obviously, there is no way currently to get the single objective tree with existing species [4], because of the evolutionary speciation events [5-8] and genes

fundamental but not totally sufficent "bearing alone" roles in species evolution history. Practically, the prediction should be compatible with job purpose and the methodology should be self-adapting.

The methodology compatible with PSC includes three steps: 1. homology assessment to get an data matrix; 2. inferring species tree by genes trees' topology; 3. methodological self-adapting [9]. We can not deal all three steps simultaneously because of computational limitation. Thus, the so called "majority enough" corresponding to the first step claims for enough data source and sufficient number of homolog genes for concatenating, considering biological phenomenon such as Lateral Gene Transfer (LGT), in order to represent species genealogy. Unfortunately, current computational methods in step one using nuclear genome or transcriptome data is inadequate to fulfil the "majority criteria" [10,11,12,13] finding the position of Xenacoelomorpha in the Metazoans concordance tree, regardless of a hundred or a thousand of genes concatenated, the alignment matrix is far less than the propotion of LGT alone in human nuclear genome (8% of total nuclear genes) [7]. Let alone other non-phylogenetic signals existing and the never proven "common history" of most nuclear genes [13].

Nevertheless, the mitochondrial genome as data source meets the "majority criteria" [1] using present alignment method though not 100% of genes in this kind of plasmon being included. Furthermore, with particularly uniparental reproduction in Metazoan [14], independent adaptive evolution rates [15], conservative biosynthesis function been selected by environment, physical segregated from nuclear genome [16], mitochondrial genes genealogy shall represent the genealogy of related species at a certain extent, which can be depicted as 'primary concordance tree' of Metazoan, to position Xenacoelomorpha.

The position of Xenacoelomorpha in the tree of life remains a unresolved question

Morphological studies previously suggest Xenoturbella Bocki as a turbellarian flatworm in 1949 and in 1997 position Xenoturbella either as a basal bilateria or with molluscan [17-21]. we don't discuss their result in this article as PSC we chosen.

Molecular studies based primarily on nuclear genes place Xenoturbellida as sistergroup of the Ambulacraria (echinoderms + hemichordates) [2,3,12], a basal Deuterostomes [12,23], or sistergroup to all other bilateria [10,11,24]. Using mitochondrial sequences prefer to support a basal Deuterostome placement [12,26] .

We briefly summarise their reasult and find out that the prediction based on nuclear genes can not congruent with each other. Although some problem has been resolved, such as sequence contaminated [27,28], rRNA compositional bias [10,11,29], or topolgy misleading by inadequate choice of other Metazoan clades [10,11,12,22] but their limited alignment matrix are not majority enough to represent nuclear genome as we discussed above. Conversely, most mitochondiral genome analyses leads to congruent prediction but being questioned by their unrepeatable results with low statistical support. Both nuclear and mitochondria data suggest to group Xenoturbella with Acoelomorph as a clade named Xenacoelomorpha.

Phylogenomics inference from mitochondrial sequences to position Xenacoelomorpha always been complicated by several misleading factors such as compositional heterogeneity [30], Long Branch Attraction (LBA) [9] and accelerated substitution rates, which violate the models of evolution. Each of them have complex source of cause and can not been absolutely separated from each other. Unfortunately, useful approaches has been used to alleviate system error but barely satisfactory.

In 2006, Bourlat SJ aimed to study the relationships among all Deuterostome groups with mitochondrial genome, leading to a result that Xenoturbella is a sister group of Ambulacraria but can not group Cephalochordate with Echinoderm. The global topology error probably caused by lack of phylogenetic signal within 170 nuclear proteins and dissatisfactory choice of species–the misplacing of Branchiostoma sp. among different datasets analyses [22].

In 2007 Marleen Perseke expanded Bourlat SJ's dataset and using Phylobayes CAT model to alleviate compositional bias getting a position that Xenoturbella is an early branch of Deuterostome with unreliable PP support, which claims for far more suited model and improved alignment matrix either [25].

Later, Bourlat SJ carried out Bayesian analyse with mtZOA leading to a basal placement of Xenoturbella within Deuterostome but the prediction of other clades in the same tree being fiercely criticized–grouping the urochordates with the echinoderms remains enigmatic in his study [26].

In 2011 Kevin J. Peterson's team conducting 37 well chosen species, mainly because of their low sequence saturation, supporting Xenacoelomorpha as a basal Deuterostomes. His arguing about grouping Acoelomorph and Xenoturbella together is generally accepted. Meanwhile, the low support of internode between Chordata and Ambulacraria, the compositional heterogenoity of mitochondrial genome and "slightly hand made" alignments make his work unrepeatable [10-12].

In 2017, Helen E. Robertson inferring phylogenetic tree from mitochondrial genes retrieves Xenacoelomorpha as an early branch of Deuterostome but the statistical support for his prediction is week, probably caused by dissatisfactory choice of species [23].

Here, we have first illustrated the philosophical foundation why it is necessary to predicting Xenacoelomorpha's genealogy with mitochondrial genome instead of a bag of nuclear genes. Afterwards, we briefly surveyed the relevant researches related to Xenacoelomorpha's phylogenetic position. In the next section, our analysis, inheriting the achievements of predecessors, using currently more suitable alignment methods to avoid "hand made", building varies matrix with filtered species directly accounts for compositional heterogeneity, excluding outgroups to elaborate LBA, leads to a robust consensus topology which hold the view that Xenacoelomorpha flatworms are basal Deuterostome. In sharp contrast, inferring phylogenetic position of Xenacoelomorpha by nuclear datasets is hazardous with LBA menace and can not fill the gap between methodology and philosophy.

Results

The sistergroup of deuterostome. This study uses mitochondrial genome data to position Xenacoelomorpha based on Heterochaerus australis's sequence and additional four Xenoturbellida [12]. Other species are similar to Kevin J. Peterson previously chosen. As the rate of sequence evolution is high, we choose amino-acid sequence to deal with. The sequences are alignmented by program MAFFT, strategy G-INS-1, scoring matrix for amino acid sequences BLOSUM 62, refined by Gblocks, parameters in detail in online method [31]. All Bayesian phylogenetic reconstructions are conducted by software package, PhyloBayes , and the optimal tree under the popular maximum likelihood (ML) criterion is found by RAXML in this work. Consensus tree from bayesian analysis under the mtZOA model and ML analysis with mtZOA + I + PROTGAMMA, inferred from 13 proteins concatenation alignments including all 37 Metazoans sequences. Their tree topology congruent with each other and supported by much higher statistical support than almost ever been published—monophyly of Xenacoelomorpha including Xenoturbella and Acoelomorph (PP=1 and BP=100), Xenacoelomorpha as an early branch of Deuterostome (PP=0.99 and BP=90), Kevin J. Peterson's confusion about the "relationship between chordates and ambulacrarians" being resolved (PP=1 and BP=90) (Fig.1). Thus, we lead to a solid prediction of phylogenetic position of Xenacoelomorpha with auto-alignment methods avoiding "hand made", which concurs with the finding that Xenacoelomorpha is sistergroup of all remaining bilateria based on nuclear dataset.

We choose mtZOA from a varies of models by AICc and BIC criteria, which is inferred from a dataset including Lophotrochozoans, non-chordate Deuterostomes and diploblastic Metazoans [32]. As CAT under a Bayesian analyse has been proved to be better fitted than GTR model, regarding compositional heterogeneity [12], we also try CAT+GTR+Γ 12 (Fig. 2a) and CAT+GTR+Γ 4 (Fig. 2b) based on 37 species matrix. In spite of they haven't challenged mtZOA's prediction, also positioning Xenacoelomorpha as a basal Deuterostomes, even cross-validation test proving CAT+GTR+Γ has better score than mtZOA, there are some problematic nodes in the CAT prediction which is weekly supported and can not congruent with the prediction under ML analyse. Thus,we harbor the idea that mtZOA as an empirical transition probability matrix probably has replacement rates and stationary frequencies better fitted to evolutionary process of our matrix leading to a robust result.

For fear of model over-fitting, we do ML and Bayesian analysis based on varies alignments—building two matrix using MAFFT (strategy G-INS-1, scoring matrix for amino acid sequences choosing BLOSUM45 and the other with BLOSUM30), both refined by Gblocks in the same way detailing in online method. All analyses using mtZOA model based on 37 species matrix or 39 species matrix (additional Schmidtea mediterranea and Dugesia japonica to exaggerate LBA event) can carry out congruent result under both Bayesian an ML analyses placing Xenacoelomorpha as a basal branch of Deuterostome with reliable statistical support (Fig.1 and Fig. 3 a–d). In contrast, CAT model under Bayesian analyses can not lead to a consensus concerning about Xenacoelomorpha's phylogenetic position acrossing different datasets and not congruent with ML analyses (Fig. 3 e,f and Fig.2 a,b). We then take mtZOA as a better fitted model to our 37 species matrix regardless of the possibility of over-fitting and suggest the 37 species alignments with parameter BLOSUM62 remaining the least non-phylogenetic signals inferring the most robust taxonomic placement with the strongest support to predicts Xenacoelomorpha as the basally

branching at Deuterostome and illuminates relationship between chordates, ambulacrarians, effectively (Fig. 1).

Inadequate taxon selection can exaggerate the compositional bias and always lead to LBA events claiming for carefully dataset preparing. Considering the compositional heterogeneity among species, we double the Xenoturbella monstrosa sequence in step one, and doing Z-score test with randomly selected two training sets of Cross-validation used (each one amounts to ninty percent of whole alignments statistically)—its about the epistemology that compositional heterogeneity hardly inevitably affects tree topology and we wondering logical and random datasets choice shall bias the tree topology in what extent. Filtering species with less than Z-score 10 (30 species) and Z-score 5 (19 species), at least one acoel retained for obvious reason, building these two alignment matrix, modeling with the same strategy as we have already done to the 37 species matrix. All 37, 30 and 19 species datasets inferring an consensus concerning that position Xenoacoelomorpha as a basal Deuterostome under both Bayesian and ML analyse (Fig. 4 a-e), showing robust topology based on varies selections of outgroup—6 outgroups to total 37 taxons, 5 to 30 and 3 to 19. We also wondering if excluded all other Xenoacoelomorphas, with single acoel we sequenced included in our 37 species matrix, can exaggerate LBA event but fail (Fig.4 e). Our prediction of Xenoacoelomorpha genealogy seems barely been biased by compositional heterogeneity with adequate choice of taxons.

We also try "Dayhoff groups" to reduce the potential influence of compositional heterogeneity on Phylogenomics analyses of our datasets [32]. The analyse under Bayesian using CAT +GTR + Γ50 + Dayhoff6 (Fig. 5 a) based on 37 species matrix and the analyse under ML using PROTOGAMMA + I + Dayhoff inferred from 30 species (Fig. 5 d) concurrently force Acoelomorph to the base of bilateria with significantly less well supported (PP=0.71, BP=100, respectively). In spite of that the 37 species and 19 species analysed under ML with PROTOGAMMA + I + Dayhoff position Xenoturbella as a basal Deuterostome (Fig. 5 b,c), overwhelmingly, different taxon selections can not lead to a congruent prediction, with problematic nodes and much lower likelihood respectively, indicating that the Dayhoff groups cannot reflects the evolutionary pattern of the given taxonomic sample properly .

We also reanalyzed the the three matrix (37, 30 and 19 taxon selections) after removing all of the outgroup taxa to rule out the possibility that the Xenoacoelomorpha branches at the basal of Deuterostome owing to long branch attraction. All three analyses congruent to position Xenoacoelomorpha as a basal Deuterostomes (Fig. 6 a-c BP=99,92,91) and the other Metazoan clades being slightly infected. The decipher is obvious that once we simply cut of the outgroup clades in figure 1, it will lead to the topology of Figure 6 a, standing to reason the topology being pretty similar globally suggesting that LBA was not the cause of this prediction.

Above all, the use of mtZOA which reflects the evolutionary pattern of our 37 species matrix, results in a much higher likelihood than CAT and Dayhoff groups. Phylogenetic information content of the matrix after auto-alignmenting and trimming merely violate the model of protein evolution. Analyses based on

varies kinds of taxon selections indicate a robust backbone tree of Metazoan inferring Xenoacoelomorpha as an early branch of Deuterostome.

Discussion

Majority enough

Homology or particularly orthology are such kind of congruent character which allow tree thinker to unite practical and conceptual advantages [13]. We want to figure out the species tree of life with PSC concept, underlying this concept is the epistemology—gene as the major inherited information carrier bearing the weight of most kind of molecular evolutionary events and being selected by environment all way through genealogy history. When we inertially start from "homolog gene" we shall not negelect the ture noumenon is either nuclear acid or amino acid "whole sequence" itself. We acknowledged different genes go through distinct evolution events, when we talk about gene level in PSC, we are trying to represent the species genealogy with "the majority of whole sequence of related species genome but not just " a bag of genes" represent themselves suffered from varies kinds of evolutionary selection. If not ,why should we use the interpretation that single gene tree inferred from cox1, getting such terrible statistical support because of lacking phylogenetic signal? (Additional File 1) If not, does it means that over ninety persent of Xenacoelomorpha's nuclear genes have nothing to do with the species evolution? It is commonly acknowledged that ortholog through all Metazoan lineages is scarce [34]. Apparently, current concept of homology is inadequate to fulfilling a prediction based on nuclear dataset.

Related works

Gene map of 10 Xenacoelomorpha species shows significant divergence among Xenoacoelomorpha (Additional file 2) and we see no directly reasoning between gene order and PSC philosophical foundation [24,26,26]. Heterochaerus australis's mitochondrial genome tRNA secondary structure predicted by mitos is showed in additional file 3 and the overview of Heterochaerus australis's mitochondrial genome sequences in additional 4. Analyse based on rRNA has been criticized for their compositional bias [1,14] . The analyses based on nuclear datasets haven't congruent with each other till today [10,11,12], while ignoring PSC majority criteria, can only be used to predict some nuclear genes genealogy. We resolved the queries about "hand made" and "problematic nodes" predecessors encountered, getting a strongly supported prediction and barely biased by compositional heterogeneity (Fig 1). In the mean time current concept of species and definition of homolog of genes leave us no other choice but take a robust mitochondiral genome analyse as the backbone tree of Metazoan including Xenacoelomorpha. As we all acknowledged that there is no single objective tree of life [4] and with the concept of homolog gene progressed, the position of Xenacoelomorpha will back to an open area taking advantages of nuclear genome data. Last but not least, the prediction we made has no direct relevance to the view of independent adaptation of regeneration [35]—a hypothesis about the marvelous function of Xenacoelomorpha under knife [36]. Because most of EGR (early growth response) motifs [37] response to

Egr (a putative pioneer factors of regeneration of Acoelomorph) are within the framework of nuclear genes prediction which shall be studied one gene a time and whether a model system regeneration regulation network can be a reference to inducing it in human is an independent job on species relationships.

Conclusions

Xenacoelomorpha flatworms are basal Deuterostome

All in all, the prediction of Xenoacoelomorpha's postion in Metazoan tree based on our auto–alignment dataset is solid (Fig. 1), not even the monoclade of Xenoacoelomorpha but also its within Deuterostome. The prediction is verified by analysis with CAT model and Dayholf groups acounting for compositional heterogeneity, and can not be challenged by cutting outgroups which always exaggerate LBA effect. Meanwhile, to illustrate our 37 species taxon selection is adequate enough to do the prediction and less biased by artificial species chosen, other two datasets included 30, 19 species were constructed, based on Z-score test to alleviate compositional bias and their modeling results not conflicting with each other. Reducing species included in analysis do show more and more LBA effect violating our model but do not misleading our prediction. We even relaxed the restriction of alignment criteria (MAFFT, strategy G–ins–1, BLOSUM 62, 45, 30) to build additional two datasets corresponding to each taxon selection mentioned above, introducing potential misleading signal in analysis to illustrate our 37 species auto-alignmented datasets used for prediction (used in figure 1) not an artificially restricted one and all the datasets discussed above lead to a congruent prediction which indicating that even misleading signal really exist in our dataset, it can not bias the prediction.

Methods

Getting mitochondrial sequence

Adult specimens of Heterochaerus australis been cultured and asexual reproduced in laboratory. DNA was isolated from adult specimens after starved two weeks totally shaded to avoid contamination by their food and symbiosis algae, extracted total DNA as described in A Arimoto et al. [38]. Sequencing, assembling and annotation are carried out as described in Felix Guerrero et al.[39].

Phylogenetic analysis

Concatenated amino acid alignments of twelve protein-coding gene from 39 species [12] including our Heterochaerus australis are used to infer all phylogenetic analyses in this article. Alignments are infferred by MAFFT V7 (G-ins-we, Blosom 62) and trimmed by Gblocks 0.91b (b1= default, b2= default, b3= default, b4=5, b5=h.). Maximum likelihood analyses and bootstrap test carried out by RAXML V8.2 ML + BP online platform. Bayesian analyses are executed by PHYLOBAYES V3.2.6. Cross-validatoin and posterior probably test are conducted by embedded package in PHYLOBAYES. Each Bayesian analyses carries out two separated chain and stops according to maximum difference less than 0.1.criteria.

Abbreviations

BSC: biological species concepts; PSC: Phylogenetic Species Concept; LGT: Lateral Gene Transfer; LBA: Long Branch Attraction; ML: maximum likelihood; PP: posterior probability; BP: bootstrap.

Declarations

Authors' contributions

Yi Wang conceived the project, designed the experiment.

Availability of data and materials

The GenBank accession number of Heterochaerus australis's mitochondrial genome is MN786871. Other mitochondrial genome sequences of 39 species of Metazoan from GenBank using in this study detailed in Additional File 5.

Ethics approval and consent to participate

Written informed consent for publication was obtained from all participants and ethics approval is not applicable.

Consent for publication

Written informed consent for publication was obtained from all participants.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Marine Life, Ocean University of China at Shandong,

266003, China;

References

1. Velasco JD: Species, Genes, and the tree of life. Br J Philos Sci. 2010; 61: 599-619.

2. Quentin D: Wheeler and Rudolf Meier, eds.Species concepts and phylogenetic theory: a debate. Columbia University Press. 2000; 17–29.

3. Cracraft J, Johnston RF: Species concepts and speciation analysis. Springer. 1983; 159–87.

4. LaPorte J: Is there a single objective, evolutionary tree of life? J Philos. 2005; 102: 357-374.

5. Bannert N, Kurth R: Retroelements and the human genome: New perspectives on an old relation. Proc Natl Acad Sci U S A. 2004; 101: 14572-14579.

6. De Jong P: Initial sequencing and analysis of the human genome (vol 409, pg 860, 2001). Nature. 2001; 412: 565-566.

7. Bapteste E, Burian RM: On the need for integrative phylogenomics, and some steps toward its creation. Biology & Philosophy. 2010; 25:711-736.

8. Bannert N, Kurth R: Retroelements and the human genome: New perspectives on an old relation. Proc Natl Acad Sci U S A. 2004; 101: 14572-14579.

9. Gouy R, Baurain D, Philippe H: Rooting the tree of life: the phylogenetic jury is still out. PHILOS T R SOC B. 2015; 370.

10. Rouse GW, Wilson NG, Carvajal JI, Vrijenhoek RC: New deep-sea species of Xenoturbella and the position of Xenacoelomorpha. Nature. 2016; 530: 94-+.

11. Cannon JT: Xenacoelomorpha is the sister group to Nephrozoa. Nature 2016; 530: 89-+.

12. Philippe H: Acoelomorph flatworms are Deuterostomes related to Xenoturbella. Nature. 2011; 470: 255-+.

13. Bapteste E, Susko E: Do orthologous gene phylogenies really support tree-thinking? BMC Evol Biol. 2005; 5.

14. Radzvilavicius AL, Kokko H, Christie JR: Mitigating mitochondrial genome erosion without recombination. Genetics. 2017; 207: 1079-1088.

15. James JE, Piganeau G, Eyre-Walker A: The rate of adaptive evolution in animal mitochondria. Mol Ecol. 2016; 25: 67-78.

16. Kaniak-Golik A, Skoneczna A: Mitochondria-nucleus network for genome stability. Free Radic Biol Med. 2015; 82: 73-104.

17. Westblad E: Xenoturbella bocki n.g., n. sp. a peculiar, primitive Turbellarian type. Ark Zool. 1949; 1:11-29.

18. Hejnol A: Assessing the root of bilaterian animals with scalable phylogenomic methods. Proc Biol Sci. 2009; 276(1677):4261-70.

19. Srivastava M: Whole-body acoel regeneration is controlled by Wnt and Bmp-Admp signaling. Curr Biol.2014; 24(10):1107-13.

20. Ehlers U, Sopott Ehlers B: Ultrastructure of the subepidermal musculature of Xenoturbella bocki, the adelphotaxon of the Bilateria. Zoomorphology. 1997; 117:71-79.

21. Israelsson O: ...and molluscan embryogenesis. Nature. 1997; 390(6655):32-32.

22. Bourlat SJ, Juliusdottir T: Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. Nature. 2006; 444(7115):85-88.

23. Robertson HE The mitochondrial genomes of the acoelomorph worms Paratomella rubra, Isodiametra pulchra and Archaphanostoma ylvae. Sci Rep. 2017; 7(1):1847.

24. Telford MJ: Combined large and small subunit ribosomal RNA phylogenies support a a basal position of the acoelomorph flatworms. Proc Biol Sci.2003; 270(1519):1077-83.

25. Perseke M, Hankeln T, Weich B, Fritzsch G, Stadler PF, Israelsson O, Bernhard D, Schlegel M: The mitochondrial DNA of Xenoturbella bocki : genomic architecture and phylogenetic analysis. Theory Biosci. 2007, 126(1):35-42.

26. Bourlat SJ: The mitochondrial genome structure of Xenoturbella bocki (phylum Xenoturbellida) is ancestral within the Deuterostomes. BMC Evol Biol. 2009; 9:107

27. Noren M; Jondelius U: Xenoturbellas molluscan relatives. Nature. 1997; 390:31-32.

28. Bourlat SJ: Xenoturbella is a Deuterostome that eats molluscs. Nature. 2003; 424:925-928.

29. Nesnidal MP, Helmkampf M, Bruchhaus I, Hausdorf B: Compositional heterogeneity and phylogenomic inference of metazoan relationships. Mol Biol Evol. 2010; 27: 2095-2104.

30. Foster PG: Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. J Mol Evol. 1999; 48(3):284-90.

31. Nakamura T, Yamada KD, Tomii K, Katoh K: Parallelization of MAFFT for large-scale multiple sequence alignments. Bioinformatics. 2018; 34: 2490-2492.

32. Rota-Stabelli O, Yang Z, Telford MJ: mtZOA: A general mitochondrial amino acid substitutions model for animal evolutionary studies. Mol Phylogenet Evol. 2009; 52: 268-272.

33. Nesnidal MP, Helmkampf M, Bruchhaus I, Hausdorf B: Compositional Heterogeneity and Phylogenomic Inference of Metazoan Relationships. Mol Biol Evol. 2010; 27: 2095-2104.

34. Lartillot N, Brinkmann H, Philippe, H: Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. BMC Evol Biol. 2007; 7.

35. Maden M: The evolution of regeneration – where does that leave mammals? Int J Dev Biol.2018; 62(6-7-8):369-372

36. Egger B: Free-living flatworms under the knife: past and present. Dev Genes Evol.2007; 217(2):89-104.

37. Gehrke AR: Acoel genome reveals the regulatory landscape of whole-body regeneration. 2019; 363(6432).

38. Arimoto A, Hikosaka-Katayama T: A draft nuclear-genome assembly of the acoel flatworm Praesagittifera naikaiensis. GigaScience. 2019; 8: issue 4.

39. Ramakodi MP, Singh B, Wells JD, Guerrero F, Ray DA: A 454 sequencing approach to dipteran mitochondrial genome research. Genomics. 2015; 105: 53-60.
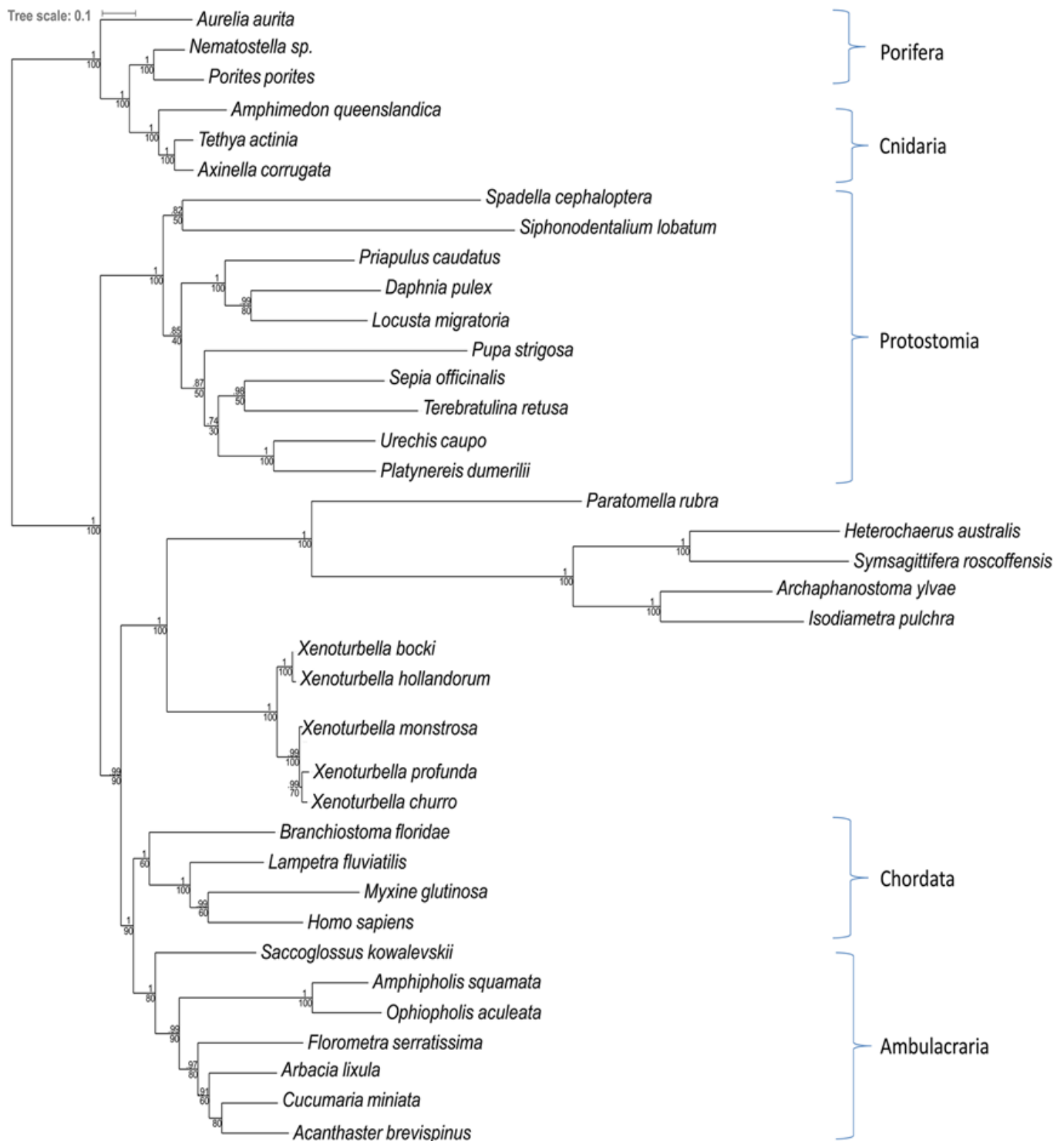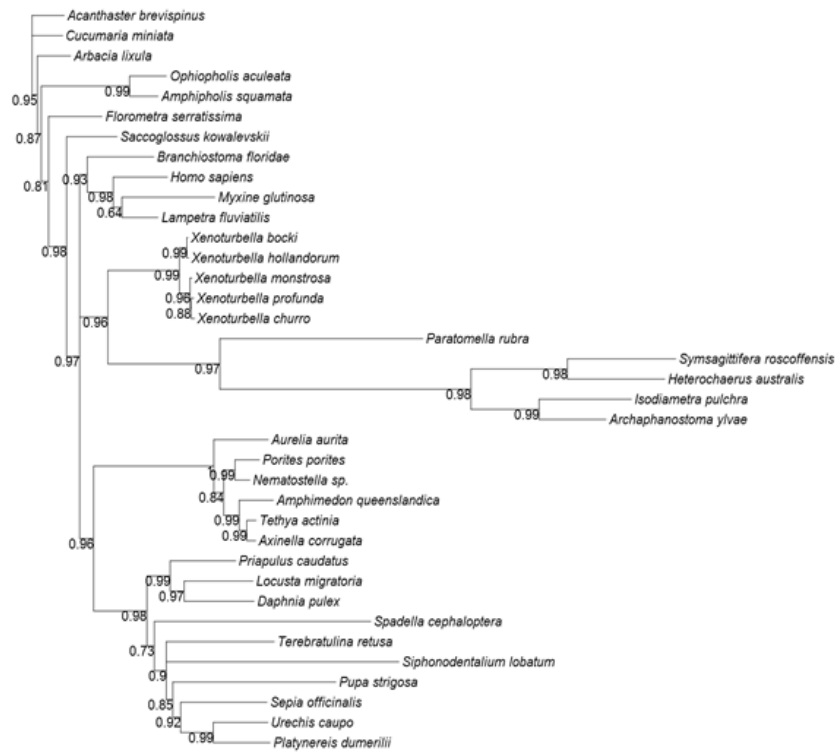
Figures

Figure 1

Xenacoelomorpha as a basal Deuterostome. Animal phylogeny based on mitochondrial proteins reconstructed using mtZOA under a Bayesian analyse and mtZOA + I + PROTGAMMA under a Maximum likelihood analyse. Xenoturbella and Acoelomorph consititute a monophyly named Xenacoelomorpha (PP=1 and BP=100), Xenacoelomorpha as a basal Deuterostome strongly supported (PP=0.99 and BP=90), placing chordates as a sister group of ambulacrarians (PP=1and BP=90) and recovering the

monophyly of Deuterostomia (PP=0.99 and BP=90). Topology of both trees is nearly identical. Analysis carried out on auto-trimmed alignments. Scale units are substitutions per site.
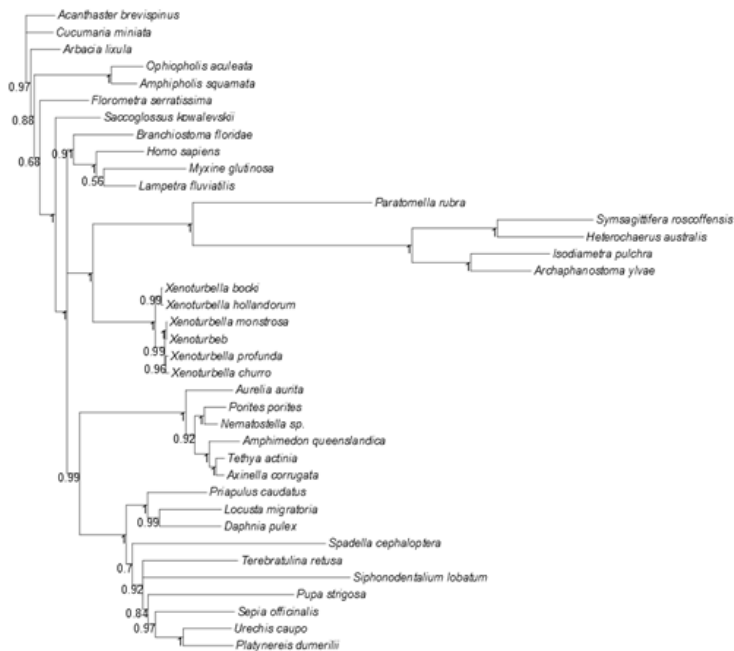

Figure 2

Choosing CAT acounts for compositional heterogeneity. (a) Phylogeny of 37 species based on mitochondrial proteins reconstructed using CAT+GTR+Γ12 under a Bayesian analyse, places Xenacoelomorpha as a sistergroup of chordates with dissatisfactory internode support. (b) Phylogeny of

37 species based on mitochondrial proteins reconstructed inferred with CAT + GTR + Γ4 in a Bayesian analyse, problematically groups Xenacoelomorpha with chordates missing node support either. Cross-validation demonstrates that CAT + GTR + Γ is better fitted model comparing to mtZOA with significantly statistical support (ΔlnL=94±14), but using mtZOA can lead to a prediction strongly supported globally. As a matter of fact, missing internode support can prove nothing but Xenacoelomorpha can be seen as a clade of Deuterostomes (PP=0.97 and PP = 1, respectively. "Xenoturbeb" is "another Xenoturbella monstrosa" illustrated in following section).
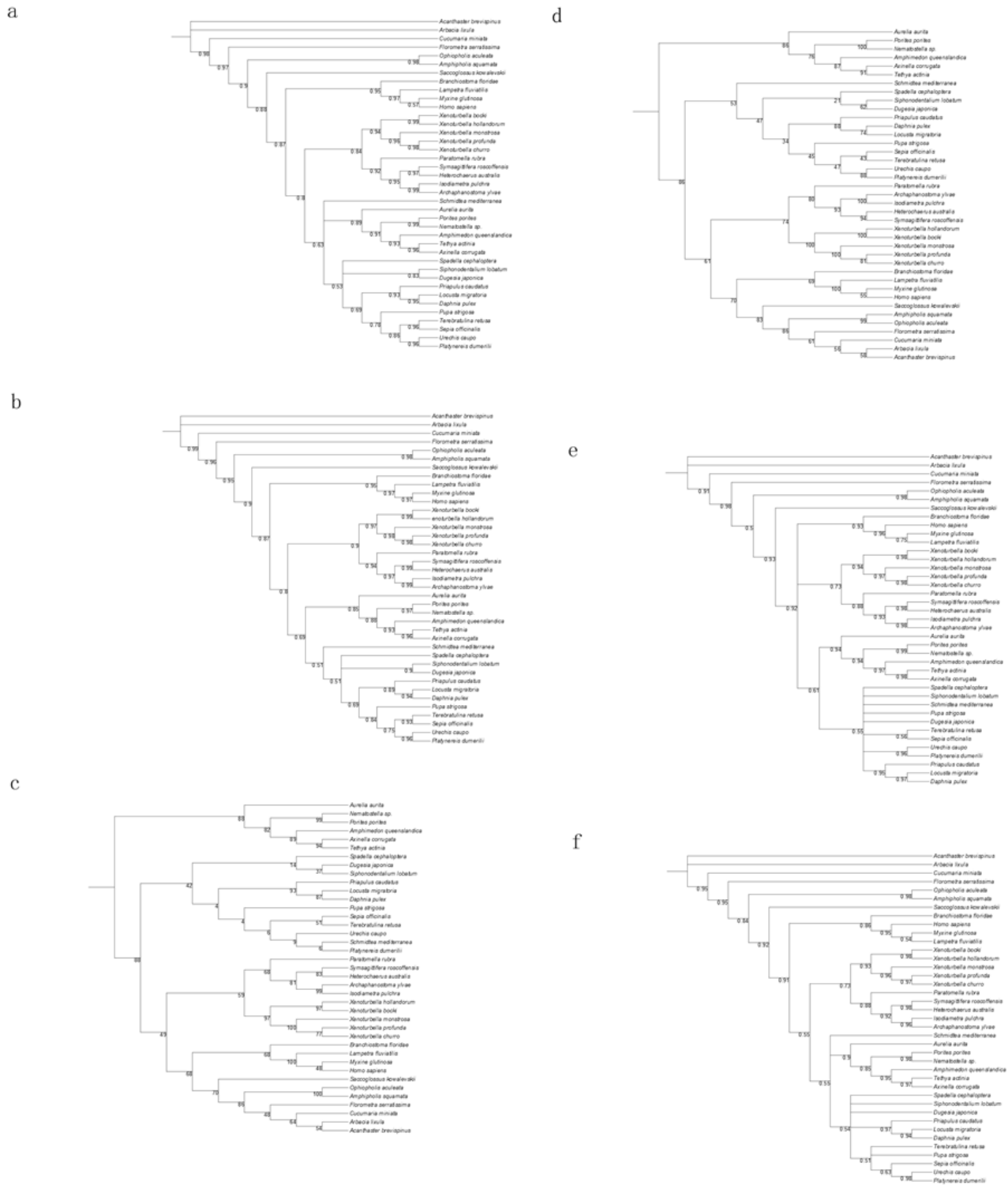
Figure 3

Comparing mtZOA and CAT model with varies aliment strategies. (a,b) Bayesian analyse using mtZOA based on matrixes conducted by alignment parameter BLOSUM30 and BLOSUM45 as the scoring matrix for amino acid sequences, position Xenacoelomorpha as the most basally branching at Deuterostome, respectively (PP=0.8 and PP=0.8). (c,d) ML analyse using mtZOA + I + PROTGAMMA based on matrixes conducted by alignment parameter BLOSUM30 and BLOSUM45, predict Xenacoelomorpha as the sistergroup of all other Deuterostome with weekly supported (BP=88 and BP=86), respectively. (e) Bayesian analyse using CAT+GTR+Γ4 based on matrixes conducted by alignment parameter BLOSUM30, position Xenacoelomorpha as a problematic clade of chordates, missing PP support. (f) Bayesian analyse using CAT+GTR+Γ4 based on matrixes conducted by alignment parameter BLOSUM45, position Xenacoelomorpha as a basal Deuterostome (PP=0.55). (a−f) Monophyly of Xenacoelomorpha haven't been challenged but the support is dissatisfactory (PP=0.84, PP=0.9, BP=59, BP=74, PP=0.73 and PP=0.73). The most enigmatic placing of Schmidtea mediterranea who is anatomically similar to almost all Xenacoelomorpha, probably caused by LBA and show no direct infection to the taxonomic placement of Xenacoelomorpha in our molecular analyse in sharp contrast to other morphological analysis.
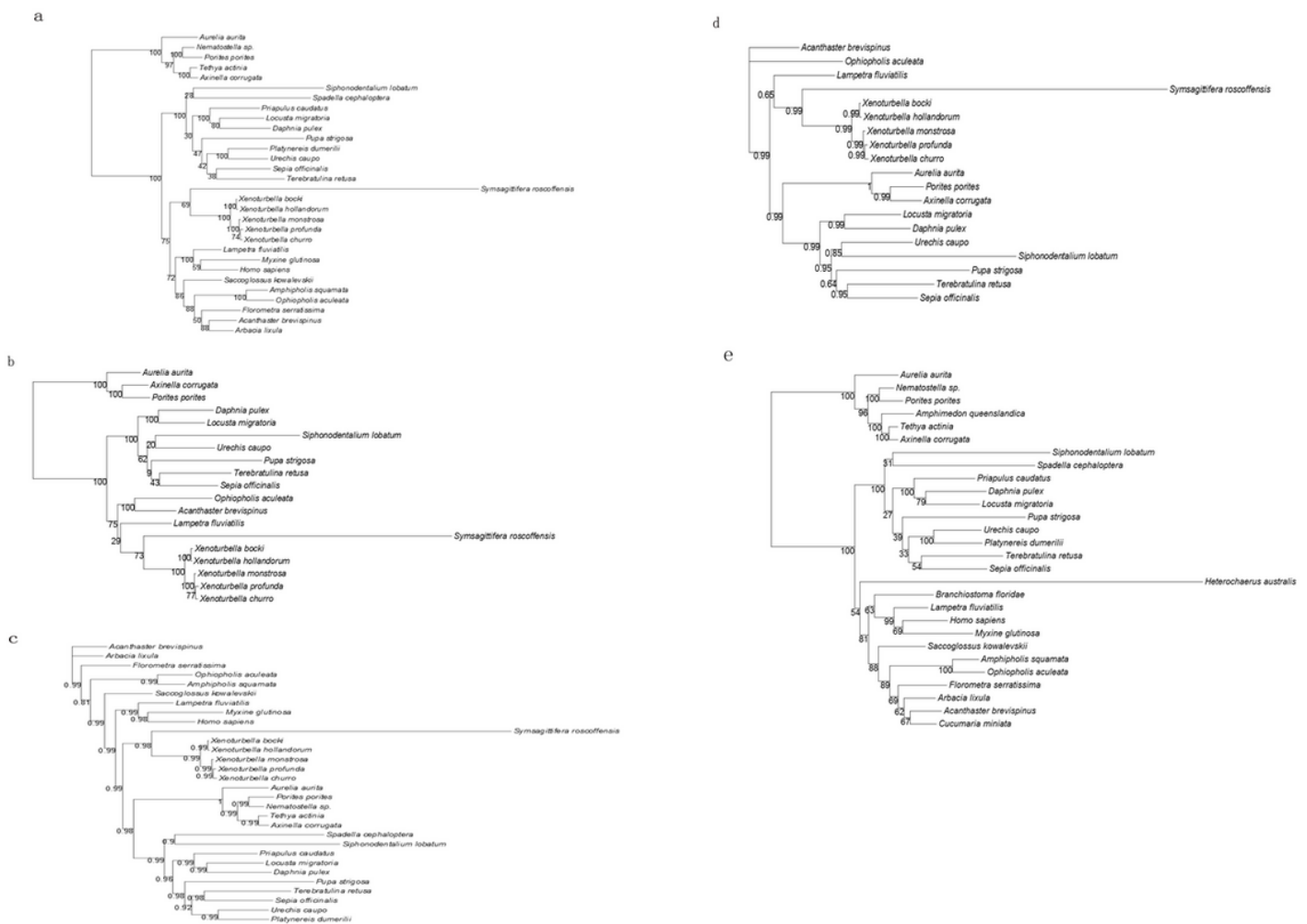

Figure 4

Varies species selection to exaggerate the violation of models. (a) 30 species phylogeny based on mitochondrial proteins reconstructed under a Maximum likelihood analysis with mtZOA + I + PROTOGAMMA, Xenacoelomorpha and five acoel species group together (BP=69) and Xenacoelomorpha within Deuterostome (BP=75) (b) 19 species phylogeny based on mitochondrial proteins reconstructed under a Maximum likelihood analysis using mtZOA + I + PROTOGAMMA, Xenacoelomorpha and five acoel species group together (BP=73) and Xenacoelomorpha as a monophyly within Deuterostome (BP=75). Weak support can be illustrated by lacking phylogenetic signal in less taxons and the tree topology still consensus with our 37 species results. (c) 30 species phylogeny based on mitochondrial proteins reconstructed under a Bayesian analyse using mtZOA, also groups Xenacoelomorpha and five acoel species together (PP=0.98) as an early branch of Deuterostome (PP=0.99). (d) 19 species phylogeny based on mitochondrial proteins reconstructed using mtZOA under a Bayesian analyse, Xenacoelomorpha and five acoel species group together (PP=0.99), Xenacoelomorpha branching at early Deuterostome (PP=0.99). (e) Using mtZOA + I + PROTOGAMMA under a Maximum likelihood analysis with our acoel included only, demonstrates no significant LBA event.
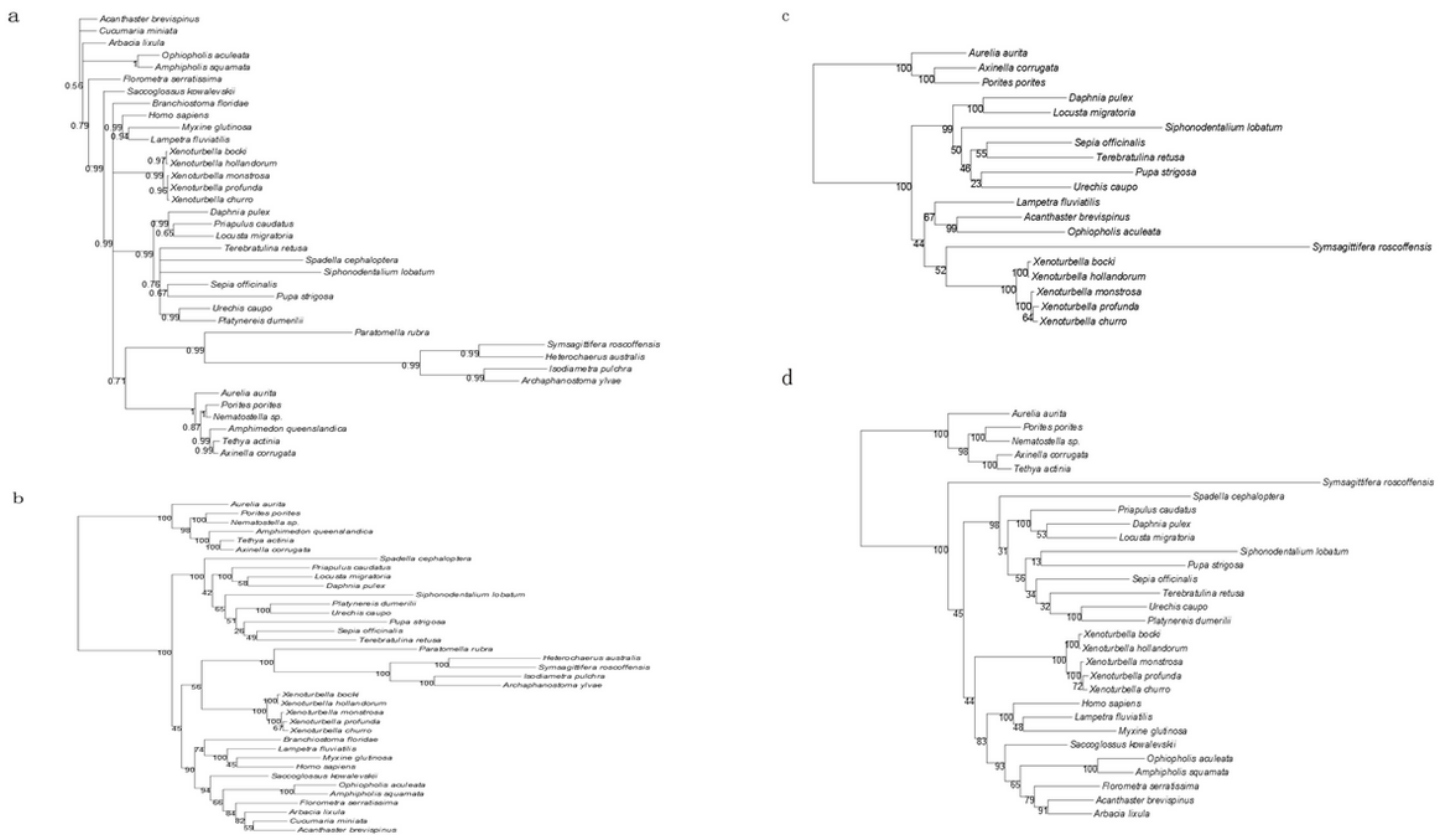

Figure 5

Reducing the potential influence of compositional heterogeneity on Phylogenomics analyses of our datasets with Dayhoff groups. (a) 37 species phylogeny based on mitochondrial proteins reconstructed using CAT +GTR + Γ50 + Dayhoff6 under a Bayesian analyse places Acoelomorph within outgroups (PP=0.71) (b) 37 species being analyzed using PROTOGAMMA + I + Dayhoff under a ML analyse, places Xenoturbella as a basal Deuterostome but with low support (BP=45) and monopoly of

Xenoacoelomorpha being challenged (BP=56). (c) 19 species analyzed under ML using PROTOGAMMA + I + dahoff , position Xenoacoelomorpha as Deuterostome (BP=44), consititutes Acoelomorph and Xenoturbella as a monophyly (BP=52). (d) The analyse of 30 species under ML, using PROTOGAMMA + I + Dayhoff, places Xenoturbella in Deuterostome (BP=44) but Acoelomorph as a basal bilateria (BP=100).
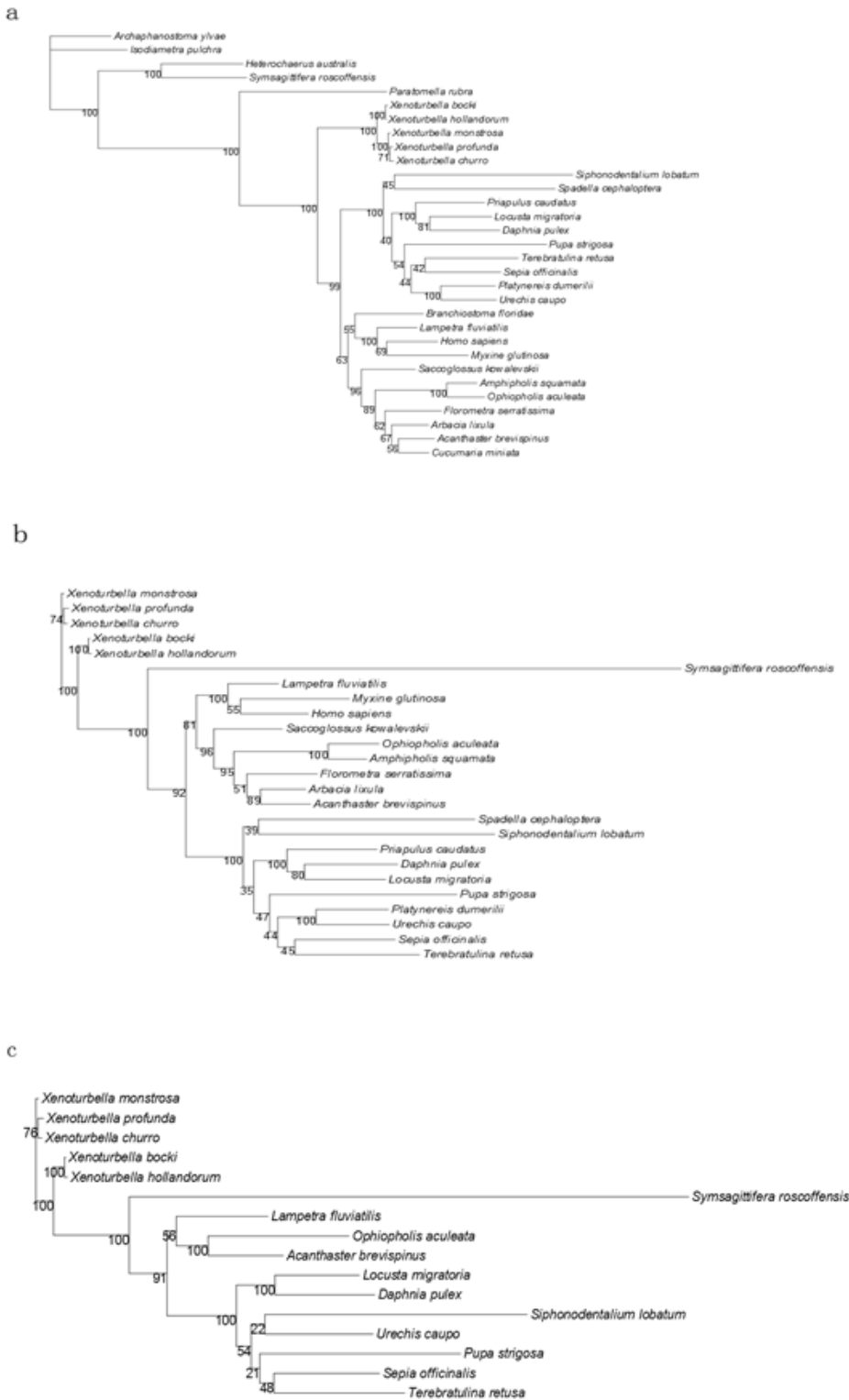

Figure 6

LBA was not the cause of this prediction. (a-c) All three taxons selections (37, 30, 19 species) without outgroups (remaining 31,25 and 16 species, respectively) position Xenoacoelomorpha as a basal Deuterostomes (BP=99,92,91) and consititution of Xenoacoelomorpha being unchallenged (BP=100,100,100), which can be comprehended as logical results after topological cutting.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1.tif
- AdditionalFile2.tif
- AdditionalFile3.tif
- AdditionalFile4.tif
- AdditionalFile5.pdf