

Supplementary Information for:

Hyoliths with pedicles constrain the origin of the
brachiopod body plan

Haijing Sun, Martin R. Smith, Han Zeng, Fangchen Zhao, Guoxiang Li and Maoyan Zhu

2018-05-28

Contents

Supplementary Text	3
1 Phylogenetic dataset	4
2 Parsimony analysis	6
2.1 Search parameters	6
2.2 Analysis	6
2.3 Results	7
3 Character reconstructions	14
3.1 Sclerites	15
3.2 Sclerites: Bivalved [2]	16
4 Fitch parsimony	20
4.1 Results	20
5 Bayesian analysis	24
5.1 Parameter estimates	25
5.2 Results	25
6 Taxonomic implications	27
Supplementary Figures	29
Supplementary Table	33
Bibliography	34

Supplementary Text

This document contains supplementary material to Sun et al. (2018). It is best viewed in HTML format at ms609.github.io/hyoliths.

It opens with a detailed discussion of analyses of the morphological dataset constructed to accompany Sun et al. (2018), and their results.

The results presented in the main paper employ the algorithm described by Brazeau et al. (2018) for correct handling of inapplicable data in a parsimony setting. This document depicts how each character is most parsimoniously reconstructed on an optimal tree.

For completeness, we also document the results of standard Fitch parsimony analysis, and the results of Bayesian analysis, neither of which treat inapplicable data in a logically consistent fashion.

Supplementary figures and tables appear after the text.

Chapter 1

Phylogenetic dataset

Analysis was performed on a new matrix of 37 early brachiozoan taxa, including hyoliths, tommotiids and mickwitziids, which were coded for 107 morphological characters (62 neomorphic, 45 transformational).

Namacalathus was incorporated as a 38th taxon, but preliminary results did not uphold the homology of its potentially brachiozoan-like features. As such, we excluded it from our analysis due to its morphological distance from ingroup taxa, a likely source of long branch error. *Dalyatia* was instead selected as an outgroup as camenellans have been interpreted as the earliest diverging members of the Brachiozoa (Skovsted et al., 2015; Zhao et al., 2017).

Characters are coded following the recommendations of Brazeau et al. (2018):

- We have employed reductive coding, using a distinct state to mark character inapplicability. Character specifications follow the structural syntax of Sereno (2007) in order to highlight ontological dependence between characters and emphasize the structure of the dataset.
- We have distinguished between neomorphic and transformational characters (*sensu* Sereno, 2007) by reserving the token 0 to refer to the absence of a neomorphic (i.e. presence/absence) character. The states of transformational characters (i.e. characters that describe a property of a feature) are represented by the tokens 1, 2, 3, ...
- We code the absence of neomorphic ontologically dependent characters (*sensu* Vogt, 2017) as absence, rather than inapplicability.

The complete dataset comprises 3959 character codings, of which 451 are inapplicable and 2319 were neither ambiguous nor inapplicable. The amount and quality of data that *is* coded is more instructive than a measure of how many cells are ambiguous (Wiens, 1998, 2003). Of the 107 characters, the number that were coded with an applicable token for each taxon is:

<u>Acanthotretella spinosa</u>	59 	<u>Haplophrentis carinatus</u>	65 	<u>Orthis</u>
<u>Alisina</u>	75 	<u>Helio medusa orienta</u>	57 	<u>Paterimitra</u>
<u>Askepasma toddense</u>	65 	<u>Kutorgina chengjiangensis</u>	76 	<u>Pedunculotheca diania</u>
<u>Antigonambonites planus</u>	73 	<u>Lingula</u>	93 	<u>Pelagodiscus atlanticus</u>
<u>Botsfordia</u>	64 	<u>Lingulosacculus</u>	49 	<u>Phoronis</u>
<u>Clupeafumosus socialis</u>	66 	<u>Lingulellotreta malongensis</u>	70 	<u>Salanygolina</u>
<u>Coolinia pecten</u>	69 	<u>Longtancunella chengjiangensis</u>	49 	<u>Siphonobolus priscus</u>
<u>Craniops</u>	57 	<u>Micrina</u>	59 	<u>Terebratulina</u>
<u>Dailyatia</u>	32 	<u>Micromitra</u>	69 	<u>Ussunia</u>
<u>Eccentrotheca</u>	31 	<u>Mickwitzia muralensis</u>	63 	<u>Tomteluva perturbata</u>
<u>Eoobolus</u>	70 	<u>Mummpikia nuda</u>	45 	<u>Yuganotheca elegans</u>
<u>Glyptoria</u>	65 	<u>Nisusia sulcata</u>	73 	
<u>Gasconsia</u>	61 	<u>Novocrania</u>	86 	

The matrix can be viewed interactively and downloaded at Morphobank (project 2800). [This link will become live on publication of the paper. Referees should follow the pre-publication link to the dataset that has been provided in the main manuscript.]

A static version of the NEXUS file used to generate this supplementary information can be downloaded directly from https://raw.githubusercontent.com/ms609/hyoliths/master/mbank_X24932_4-18-2018_656.nex.

Chapter 2

Parsimony analysis

The phylogenetic dataset contains a considerable proportion of inapplicable codings ($451/3959 = 11.4\%$ of tokens), which are known to introduce error and bias to phylogenetic reconstruction when the Fitch algorithm is employed (Maddison, 1993; Brazeau et al., 2018). As such, we employed a new tree-scoring algorithm that correctly handles inapplicable data (Brazeau et al., 2018), implemented in the *MorphyLib* C library (Brazeau et al., 2017). We employed the R package *TreeSearch* v0.1.2 (Smith, 2018) to conduct phylogenetic tree search with this algorithm.

As this is a new method, we also employed the traditional, Fitch algorithm, even though this approach is known to generate erroneous trees. The results of this analysis can be viewed in a later section.

2.1 Search parameters

Heuristic searches were conducted using the parsimony ratchet (Nixon, 1999) under equal and implied weights (Goloboff, 1997). The consensus tree presented in the main manuscript represents a strict consensus of all trees that are most parsimonious under one or more of the concavity constants (k) 2, 3, 4.5, 7, 10.5, 16 and 24, an approach that has been shown to produce higher accuracy (i.e. more nodes and quartets resolved correctly) than equal weights at any given level of precision (Smith, 2017).

2.2 Analysis

The R commands used to conduct the analysis are reproduced below. The results can most readily be replicated using the R markdown files (.Rmd) used to generate these pages.

2.2.1 Initialize and load data

```
# Load data from locally downloaded copy of MorphoBank matrix
my_data <- ReadAsPhyDat(filename)
my_data$Namacalathus <- NULL # Exclude Namacalathus
iw_data <- PrepareDataIW(my_data)
```

2.2.2 Generate starting tree

Start by quickly rearranging a neighbour-joining tree, rooted on the outgroup.

```

nj.tree <- NJTree(my_data)
rooted.tree <- EnforceOutgroup(nj.tree, outgroup)
start.tree <- TreeSearch(tree=rooted.tree, dataset=my_data, maxIter=3000,
                           EdgeSwapper=RootedNNISwap, verbosity=0)

```

2.2.3 Implied weights analysis

The position of the root does not affect tree score, so we keep it fixed (using RootedXXXSwap functions) to avoid unnecessary swaps.

```

for (k in kValues) {
  iw.tree <- IW Ratchet(start.tree, iw_data, concavity=k,
                         ratchHits = 60, searchHits=55,
                         swappers=list(RootedTBRSwap, RootedSPRSwap, RootedNNISwap),
                         verbosity=0L)
  score <- IWScore(iw.tree, iw_data, concavity=k)
  # Write a single best tree
  write.nexus(iw.tree,
              file=paste0("TreeSearch/hy_iw_k", k, "_",
                          signif(score, 5), ".nex", collapse=''))
  
  iw.consensus <- IW RatchetConsensus(iw.tree, iw_data, concavity=k,
                                         swappers=list(RootedTBRSwap, RootedNNISwap),
                                         searchHits=55,
                                         nSearch=150, verbosity=0L)
  write.nexus(iw.consensus,
              file=paste0("TreeSearch/hy_iw_k", k, "_",
                          signif(IWScore(iw.tree, iw_data, concavity=k), 5),
                          ".all.nex", collapse=''))
}

```

2.2.4 Equal weights analysis

```

ew.tree <- Ratchet(start.tree, my_data, verbosity=0L,
                     ratchHits = 25, searchHits=55, # ratchHits = 10 not enough
                     swappers=list(RootedTBRSwap, RootedSPRSwap, RootedNNISwap))
ew.consensus <- RatchetConsensus(ew.tree, my_data, nSearch=150, searchHits = 55,
                                   swappers=list(RootedTBRSwap, RootedNNISwap),
                                   verbosity=0L)
write.nexus(ew.consensus, file=paste0(collapse='', "TreeSearch/hy_ew_",
                                       Fitch(ew.tree, my_data), ".nex"))

```

2.3 Results

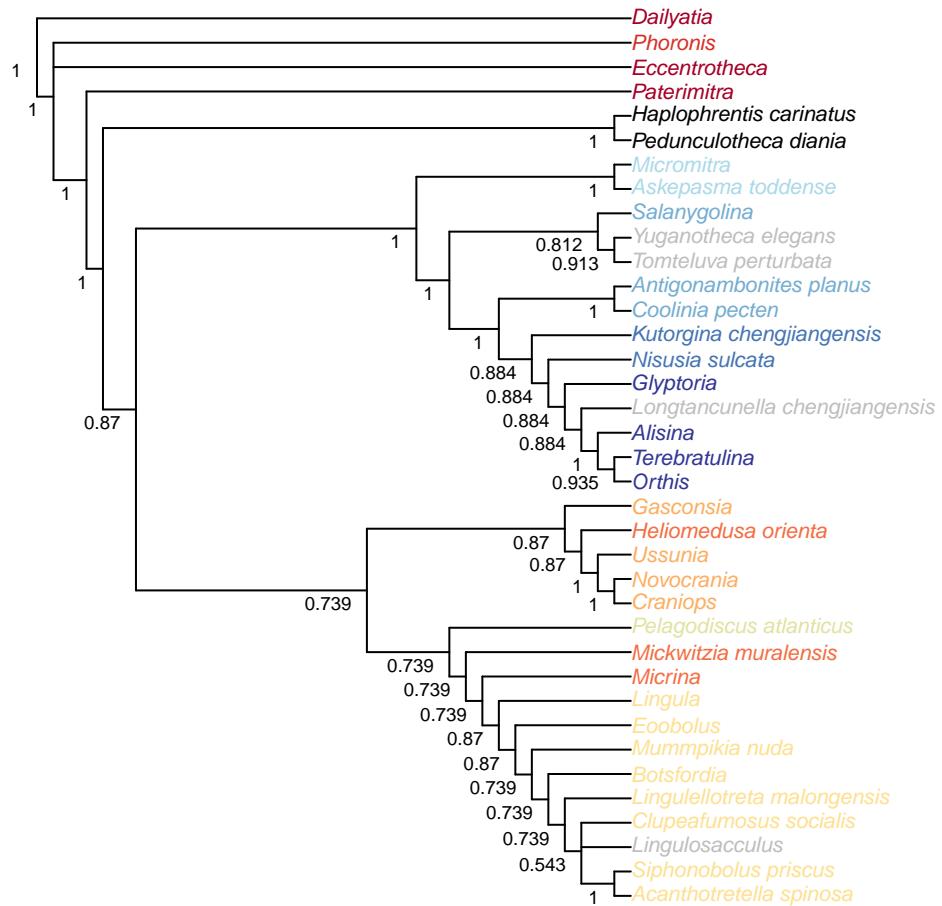


Figure 2.1: Consensus of all parsimony results. Node labels denote the proportion of trees obtained under all analytical conditions that support the clade.

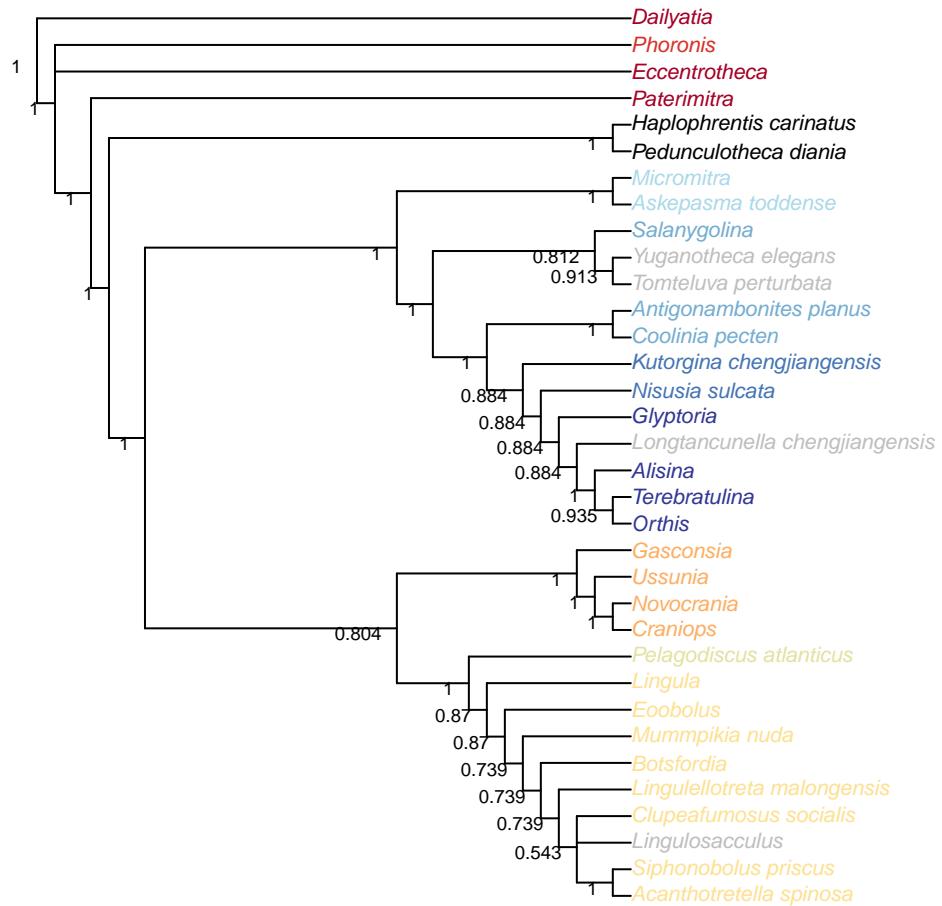


Figure 2.2: Consensus of all parsimony results. Node labels denote the proportion of trees obtained under all analytical conditions that support the clade.



Figure 2.3: Consensus of implied weights analyses at all values of k . Wildcard taxa have been excluded from the consensus tree shown above to improve resolution.

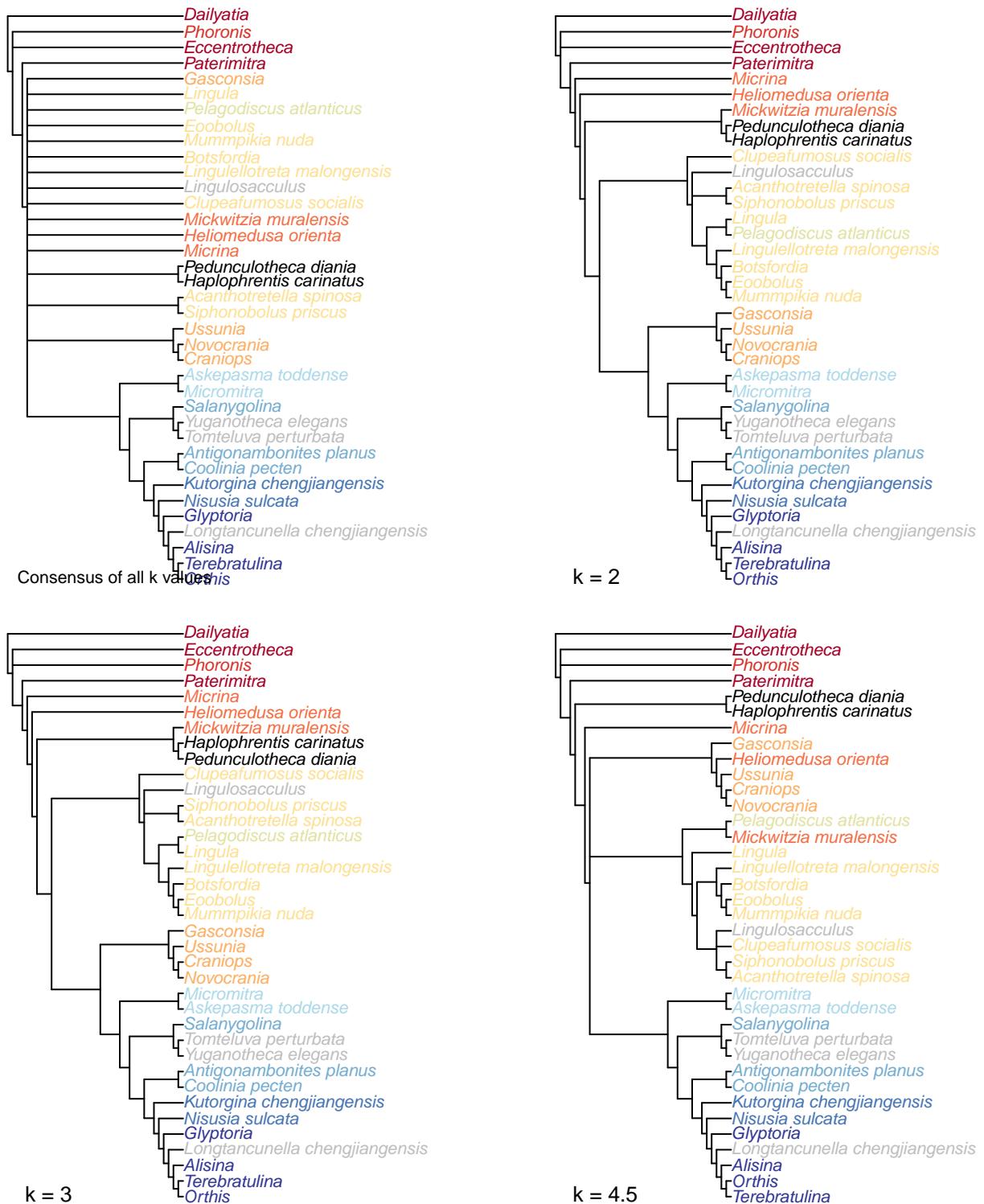


Figure 2.4: Consensus trees of implied weights analyses at all values of k , and at the individual values $k = 2, 3$ and 4.5 .



Figure 2.5: Consensus trees of implied weights analyses at $k = 7, 10.5, 16$ and 24 .



Figure 2.6: Strict consensus of most parsimonious trees under equally weighted parsimony

Chapter 3

Character reconstructions

This page provides definitions for each of the characters in our matrix, and justifies codings in particular taxa where relevant. Further citations for codings that are not discussed in the text can be viewed by browsing the morphological dataset on MorphoBank (project 2800). This link will become live on publication of the paper. Referees should follow the pre-publication link to the dataset that has been provided in the main manuscript.

Alongside each character's definition, each character is mapped onto a tree. Here, we have arbitrarily selected one most parsimonious tree obtained under implied weighting, $k = 4.5$. Other trees can be viewed in the HTML version of this file at ms609.github.io/hyoliths. Each tip is labelled as it is coded in the matrix, and these states are used to reconstruct the condition of each internal node, using the parsimony method of Brazeau et al. (2018) as implemented in the *Inapp R* package.

We emphasize that different trees will give different reconstructions. The character mappings are not intended to definitively establish how each character evolved, but to help the reader quickly establish how each character has been coded, and to visualize at a glance how well the character fits onto the given tree. We consider this more intuitive than the use of the flawed (Archie, 1989) Consistency Index, but include this value because of its historic significance.

3.1 Sclerites

[1] Present in adult



Character 1: Sclerites: Present in adult

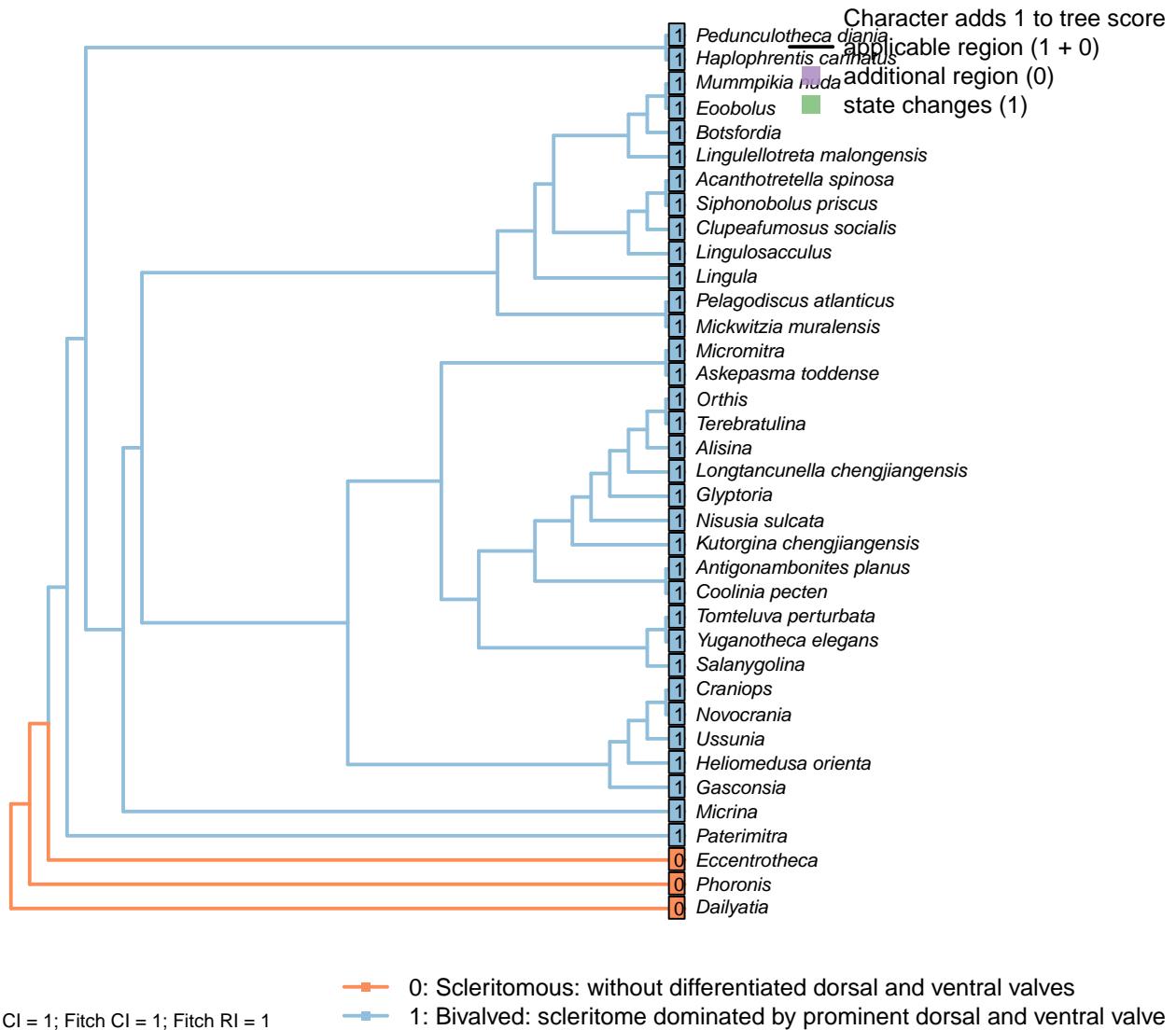
0: Absent

1: Present

Neomorphic character.

Plate-like (wider than tall) skeletal elements, whether mineralized or non-mineralized.
 The definition deliberately excludes setae (which are taller than wide).

3.2 Sclerites: Bivalved [2]

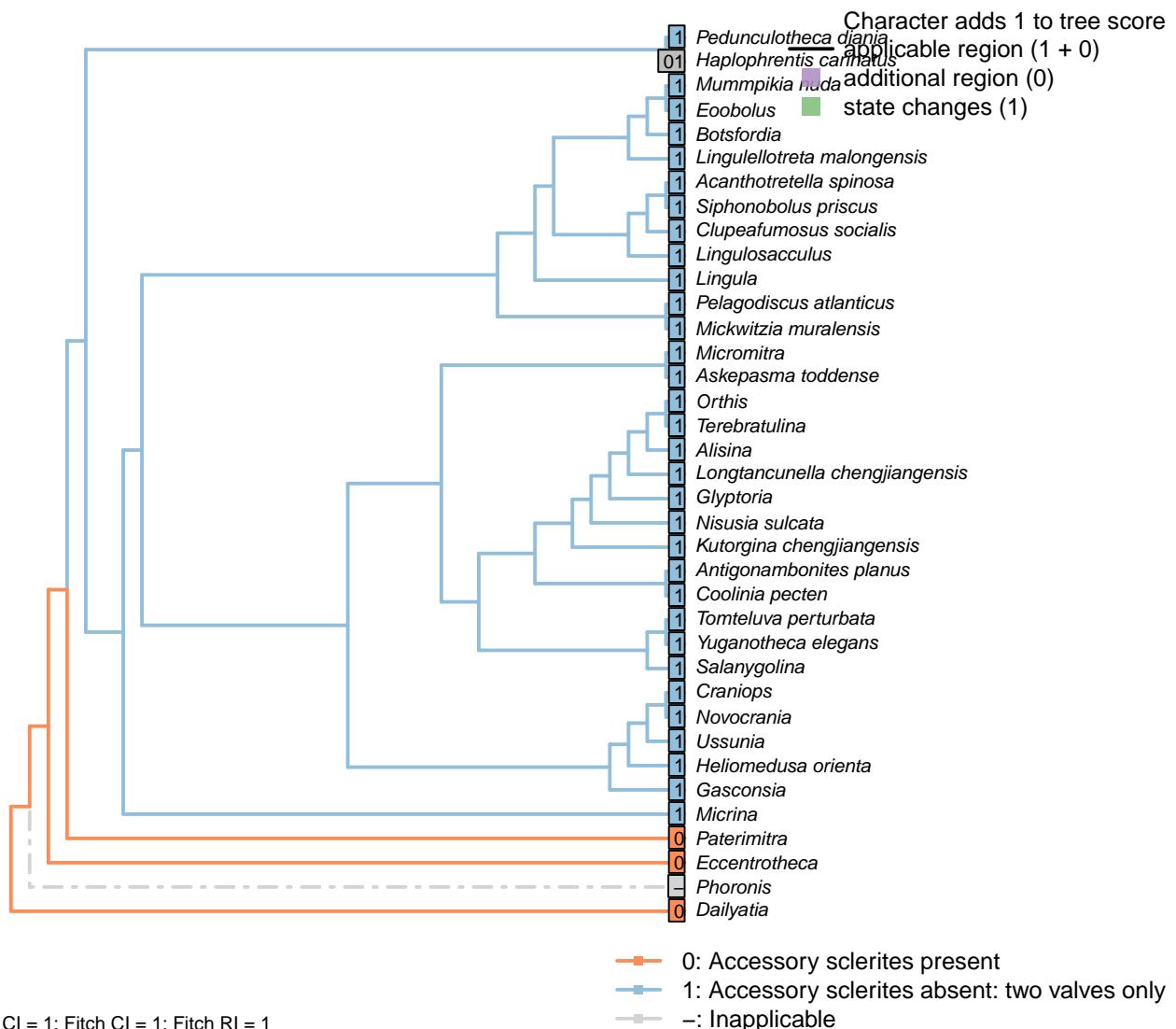


Character 2: Sclerites: Bivalved

- 0: Scleritomous: without differentiated dorsal and ventral valves
 1: Bivalved: scleritome dominated by prominent dorsal and ventral valve
 Neomorphic character.

Scleritome dominated by prominent differentiated dorsal and ventral valves.

[3] Accessory sclerites reduced



Character 3: Sclerites: Bivalved: Accessory sclerites reduced

0: Accessory sclerites present

1: Accessory sclerites absent: two valves only

Neomorphic character.

Taxa in the bivalved condition may retain sclerites as small additional elements, such as the L-elements of *Paterimitra* (Skovsted et al., 2015).

This character is treated as neomorphic, with accessory sclerites ancestrally present, recognizing the likely origin of brachiozoans (and Lophotrochozoans more generally) from a scleritomous organism.

Coded as inapplicable in taxa that lack multiple skeletal elements.

Haplophrentis carinatus: Coded as ambiguous to recognize the possibility that helens may correspond to L-elements of *Paterimitra* (Moysiuk et al., 2017).

Paterimitra: L-sclerites (Skovsted et al., 2009).

[4] Hinge line shape



Character 4: Sclerites: Bivalved: Hinge line shape

1: Astrophic

2: Strophic

Transformational character.

Botsfordia: Coded as dissociated in Williams *et al.* (1998), appendix 2.

Craniops: Astrophic: rounded posterior margin (see fig. 91 in Williams *et al.*, 2000).

Gasconsia: The straight posterior margin of *Gasconsia* contributes to an overall resemblance with the Chileids (?).

Kutorgina chengjiangensis: Williams *et al.* (2000, p. 208) consider the hinge of *Kutorgina* to be strophic, whereas Bassett *et al.* (2001) argue for an astrophic interpretation – whilst noting that the arrangement is prominently different from other astrophic taxa. We therefore code this taxon as ambiguous.

Longtancunella chengjiangensis: “*Longtancunella* has an oval to subcircular shell with a very short strophic

hinge line” – Zhang et al. (2011).

Mickwitzia muralensis: non-strophic.

Micrina: See Holmer et al. (2008).

Nisusia sulcata: “The strophic, articulated shells of the Kutorginata rotated on simple hinge mechanisms that are different from those of other rhynchonelliforms” (Williams et al., 2000, p. 208).

Novocrania: Craniides have a strophic posterior valve edge (Williams et al., 2007, table 39 on p. 2853): *Novocrania*’s “dorsal posterior margin” is “straight” (Williams et al., 2000, p. 171).

Tomteluva perturbata: “Tomteluvid taxa all have a strongly ventribiconvex, astrophic shell with a unisulcate commissure” – Streng et al. (2016), p5.

Yuganotheca elegans: Not evident from fossil material; the possibility of a short strophic hinge line (as in *Longtancunella*) is difficult to discount.

Chapter 4

Fitch parsimony

Parsimony search was conducted in TNT v1.5 (Goloboff and Catalano, 2016) using ratchet and tree drifting heuristics (Goloboff, 1999; Nixon, 1999), repeating the search until the optimal score had been hit by 1500 independent searches:

```
xmult:rat10 drift10 hits 1500 level 4 chklevel 5;
```

Searches were conducted under equal weights and results saved to file:

```
piwe-; xmult; /* Conduct search with equal weighting */  
tsav *TNT/ew.tre;sav;tsav/; /* Save results to file */  
keep 0; hold 10000; /* Clear trees from memory */
```

Further searches were conducted under extended implied weighting (Goloboff, 1997, 2014), under the concavity constants 2, 3, 4.5, 7, 10.5, 16 and 24:

```
xpiwe=; /* Enable extended implied weighting */  
piwe=2; xmult; /* Conduct analysis at k = 2 */  
tsav *TNT/xpiwe2.tre; sav; tsav/; /* Save results to file */  
keep 0; hold 10000; /* Clear trees from memory */  
piwe=3; xmult; /* Conduct analysis at k = 3 */  
tsav *TNT/xpiwe3.tre; sav ;tsav/; /* Save results to file */
```

We acknowledge the Willi Hennig Society for their sponsorship of the TNT software.

4.1 Results

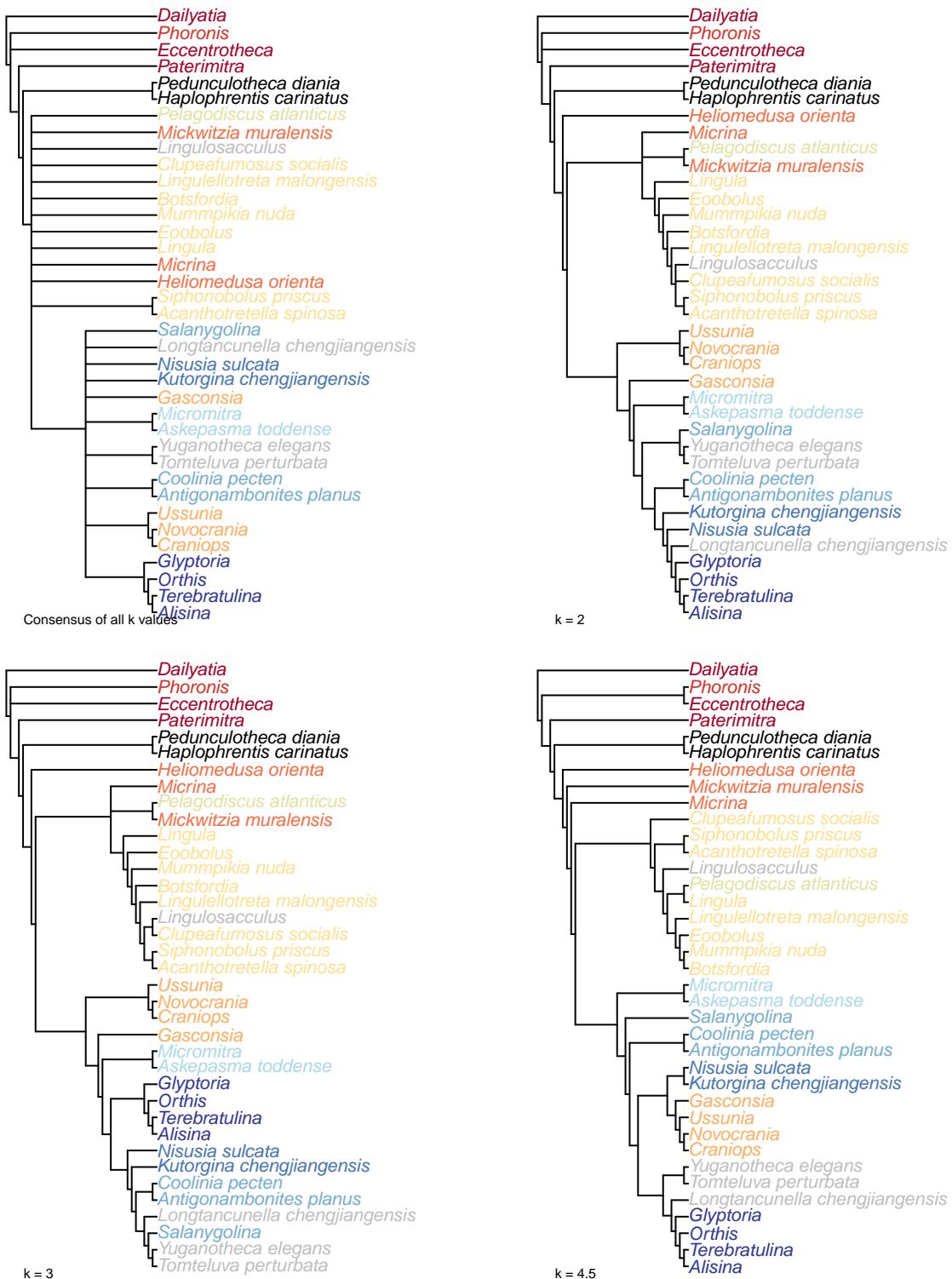


Figure 4.1: Strict consensus of all trees recovered by TNT using Fitch parsimony with implied weighting at all values of k , and at the individual values $k = 2, 3$ and 4.5 . The consensus of all implied weights runs is not very well resolved, largely due to a few wildcard taxa, particularly at $k = 4.5$, which obscures a consistent set of relationships between the remaining taxa.

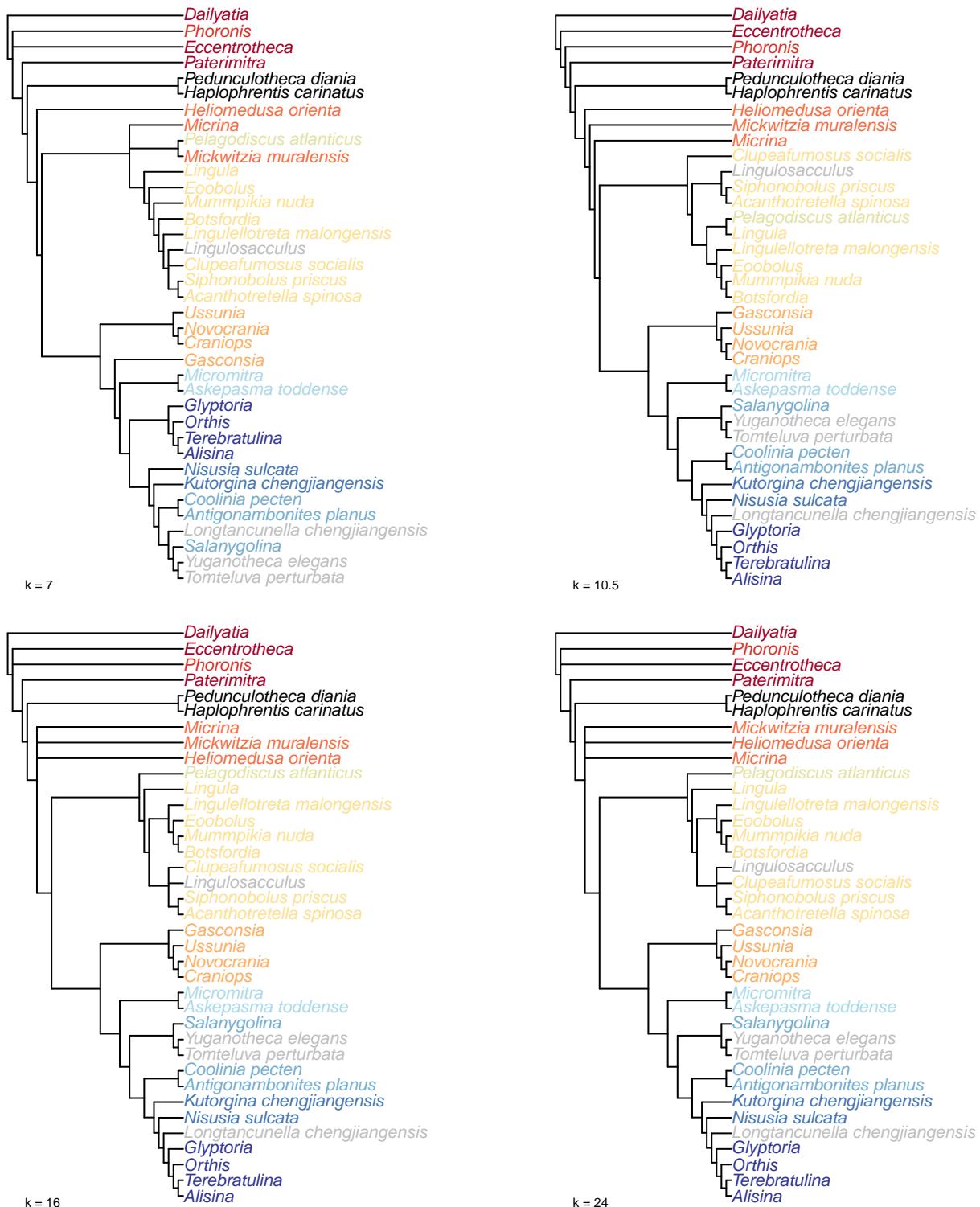


Figure 4.2: Strict consensus of all trees recovered by TNT using Fitch parsimony with implied weighting, at $k = 7, 10.5, 16$, and 24 .

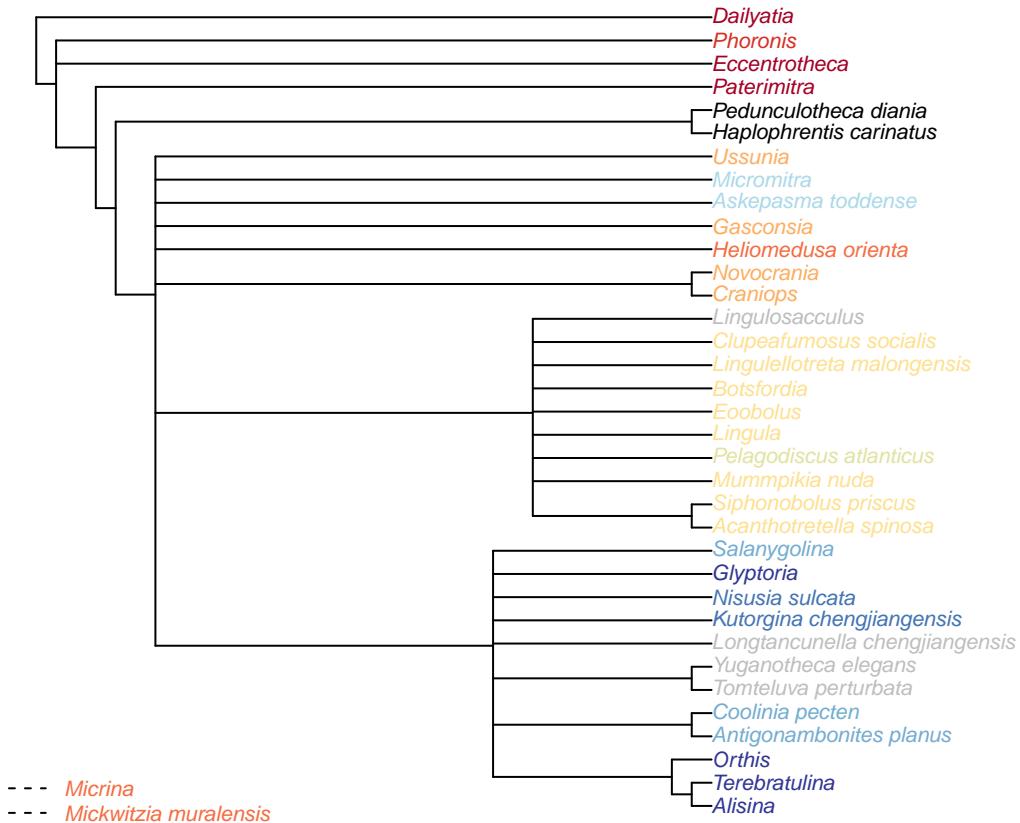


Figure 4.3: Consensus of all trees obtained using equally weighted Fitch parsimony in TNT. *Mickwitzia* and *Micrina* may equally parsimoniously be reconstructed in the basal region of the linguliform or rhynchonelliform lineages; as such, the inclusion of these taxa in the consensus tree reduces resolution. These taxa were still included in the analysis used to generate this tree, but were removed from each MPT before the consensus was calculated in order that the relationships that are present in each tree might be more easily observed.

Chapter 5

Bayesian analysis

Bayesian search was conducted in MrBayes v3.2.6 (Ronquist et al., 2012) using the Mk model (Lewis, 2001) with gamma-distributed rate variation across characters:

```
lset coding=variable rates=gamma;
```

Branch length was drawn from a dirichlet prior distribution, which is less informative than an exponential model (Rannala et al., 2012), but requires a prior mean tree length within about two orders of magnitude of the true value (Zhang et al., 2012). To satisfy this latter criterion, we specified the prior mean tree length to be equal to the length of the most parsimonious tree under equal weights, using a Dirichlet prior with $\alpha_T = 1$, $\beta_T = 1/(equal\ weights\ tree\ length/number\ of\ characters)$, $\alpha = c = 1$:

```
prset brlenspr = unconstrained: gammadir(1, 0.34, 1, 1);
```

Neomorphic and transformational characters (*sensu* Sereno, 2007) were allocated to two separate partitions whose proportion of invariant characters and gamma shape parameters were allowed to vary independently:

```
charset Neomorphic = 1 2 3 5 7 8 9 10 12 14 17 18 19 22 23 24 25 28 29 30 31 32 35 37 39 46 47  
49 50 51 52 53 54 57 61 62 63 64 66 67 69 70 71 72 73 77 78 79 81 83 86 87 91 93 94 95 97 98 99  
102 103 105;
```

```
charset Transformational = 4 6 11 13 15 16 20 21 26 27 33 34 36 38 40 41 42 43 44 45 48 55 56  
58 59 60 65 68 74 75 76 80 82 84 85 88 89 90 92 96 100 101 104 106 107;
```

```
partition chartype = 2: Neomorphic, Transformational;
```

```
set partition = chartype;
```

```
unlink shape=(all) pinvar=(all);
```

Neomorphic characters were not assumed to have a symmetrical transition rate – that is, the probability of the absent → present transition was allowed to differ from that of the present → absent transition, being drawn from a uniform prior:

```
prset applyto=(1) symdirihyperpr=fixed(1.0);
```

The rate of variation in neomorphic characters was also allowed to vary from that of transformational characters:

```
prset applyto=(1) ratepr=variable;
```

Dailyatia was selected as an outgroup:

```
outgroup Dailyatia;
```

Four MrBayes runs were executed, each sampling eight chains for 5 000 000 generations, with samples taken every 500 generations. The first 10% of samples were discarded as burn-in.

```
mcmc ngen=5000000 samplefreq=500 nruns=4 nchains=8 burninfrac=0.1;
```

A posterior tree topology was derived from the combined posterior sample of all runs. Convergence was indicated by PSRF = 1.00 and an estimated sample size of > 200 for each parameter.

5.1 Parameter estimates

Parameter	Mean	Variance	minESS	avgESS	PSRF
TL{all}	7.380	0.54000	3160	5380	0.99996
m{1}	0.609	0.00984	733	2790	1.00000

5.2 Results

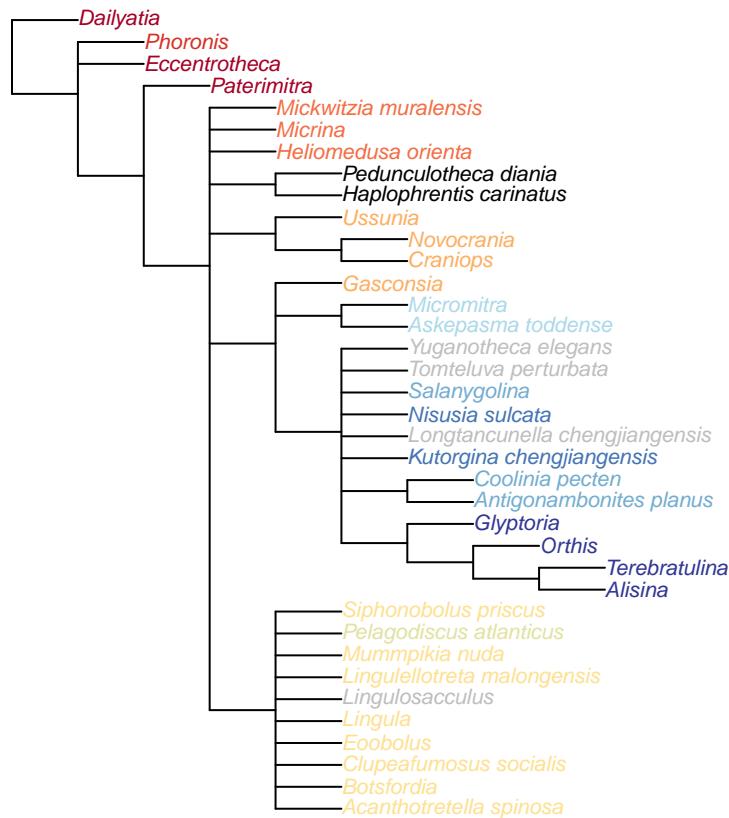


Figure 5.1: Results of Bayesian analysis, posterior probability > 50%, all taxa

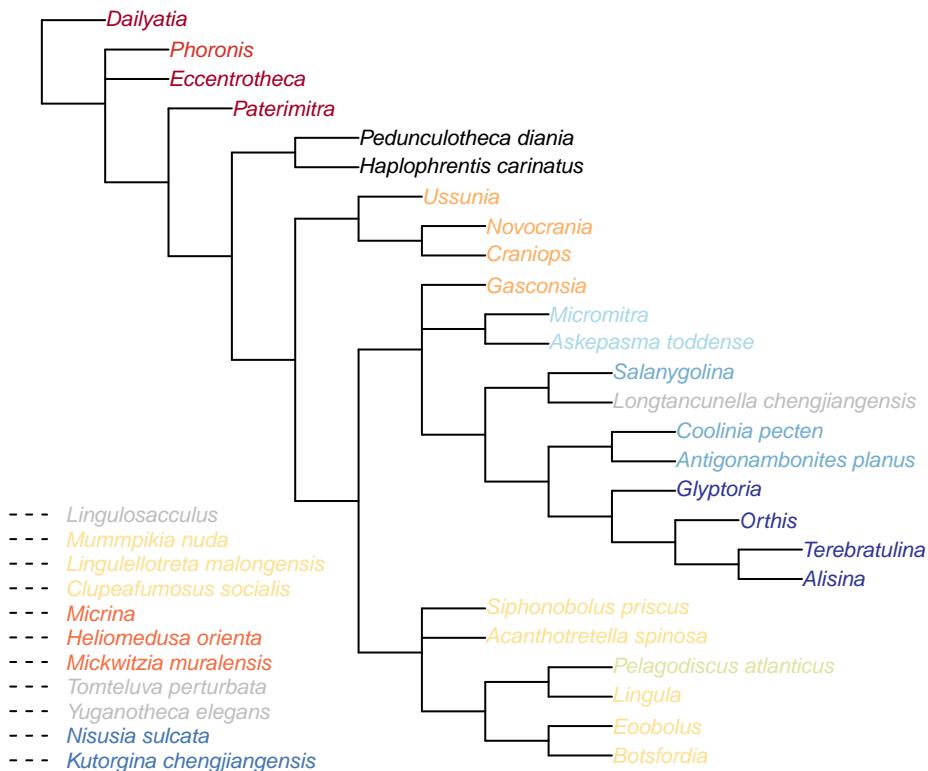


Figure 5.2: Results of Bayesian analysis, posterior probability > 50%, wildcard taxa pruned

Chapter 6

Taxonomic implications

This section briefly places key features of our results in the context of previous phylogenetic hypotheses.

Brachiopod crown and stem group Crown- and stem-group terminology has great value in clarifying the early evolution of major lineages (Budd and Jensen, 2000; Carlson and Cohen, 2009). The crown group of a lineage is defined as the last common ancestor of all living members of a group, and all its descendants; the stem group as all taxa more closely related to the crown group than to any other extant taxon. In our analyses, the brachiopod crown group corresponds to the last common ancestor of *Terebratulina* and *Lingula*; the brachiopod stem group comprises anything between this node and the branching point of *Phoronis*.

Craniiforms Trimerellids are reconstructed as paraphyletic with respect to Craniiforms. This is consistent with the affinity commonly drawn between these groups (e.g. Williams et al., 2000), and helps to account for the stratigraphically late (Ordovician) appearance of Craniids in the fossil record. (Aragonite is underrepresented in early Palaeozoic strata due to taphonomic bias.)

The relationship of Craniiforms with respect to Linguliforms and Rhynchonelliforms remains unclear. Shell characters point to a relationship with the Rhynchonelliforms, which is countered by similarities between the spermatozoa of phoronids and terebratulids, which indicate a craniiform + linguliform clade.

It's worth noting that Bayesian and Fitch analyses place *Gasconsia* as the basalmost member of the Rhynchonellid lineage, upholding suggestions (Holmer et al., 2014) of a chileid rather than trimerellid affinity. This placement presumably represents an artefact resulting from the incorrect handling of inapplicable data. But if true, *Gasconsia* would be a close analogue for the common ancestor of Rhynchonelliforms + Craniiforms (+Linguliforms?).

Rhynchonelliforms The position of kutorginids within the rhynchonelliform stem lineage has been tricky to resolve (Holmer et al., 2018b); we resolve them as paraphyletic with respect to Rhynconellata (which encompasses the obolellate *Alisina*), which is broadly in accord to previous proposals (Holmer et al., 2018a). Chileids form the adelphiotaxon to this clade. *Longtancunella* (Zhang et al., 2011) nests crownwards of the protorthid *Glyptoria*, but stemward of the obolellid *Alisina*.

Salanygolina has been interpreted as a stem-group rhynchonelliform based on its combination of paterinid and chileate features (Holmer et al., 2009). Our results position *Salanygolina* between paterinids and chileids, which directly corroborates this proposed phylogenetic position.

Basal rhynchonellids are characterized by a circular umbonal perforation in the ventral valve, associated with a colleplax. Partly on this basis, the aberrant taxa *Yuganotheca* and *Tomteluva* plot close to *Salanygolina*, the three often forming a clade – though the reliability of this grouping is perhaps liable to change as additional data comes to light. Nevertheless, an interpretation of *Yuganotheca* as a stem-group brachiopod (Zhang et al., 2014) is difficult to reconcile with the increasingly well-constrained

nature of the early brachiopod body plan.

Linguliforms The reconstruction of Linguloformea comprising Linguloidea as sister to Discinoidea is as expected, though it is notable that Acrotretids and Siphonotretids plot more closely to Linguloidea than Discinoidea does.

Lingulellotretids also sit within this lingulid grouping; a position in the phoronid stem lineage (advocated by Balthasar and Butterfield, 2009) is not upheld.

More novel is the reconstruction of the calcitic obbolellid *Mummpikia* in the linguliform total group: a rhynchonelliform affinity has been assumed based on its calcitic mineralogy. This said, Balthasar (2008) has highlighted the similarities between obbolellids and linguliform brachiopods, including sub- m vertical canals and the detailed configuration of the posterior shell margin. Our analysis upholds the case for a linguliform affinity for *Mummpikia*; a calcitic shell seemingly arose through an independent change within this taxon. As such, *Mummpikia* has no direct bearing on the origin of ‘Calciata’, save that shell mineralogy is perhaps less static than commonly assumed.

More generally, our results identify Class Obbolellata as polyphyletic: *Alisina* (Trematobolidae) plots within Rhynchonellata; *Tomteluva* is harder to place, but tends to group with *Salanygolina* stemwards of the chileids.

Paterinids Paterinids have traditionally been placed within the Linguliforms on the basis of their phosphatic shell (Williams et al., 2007), which our analysis identifies as ancestral within the brachiopod crown group; our analysis places them within the Rhynchonelliforms instead. Characters supporting this position include the strophic hinge line, planar cardinal area, the absence of a pedicle nerve impression, and the morphology of the mantle canals.

More generally, although some lingulids can be found which share more generic characters (e.g. shell growth direction) with paterinids, the particular combination of characters exhibited in paterinids does not occur anywhere in the linguliform lineage, but is more similar to that of basal rhynchonelliforms, particularly *Salanygolina*.

Tommotiids Tommotiids represent a basal grade, paraphyletic to phoronids and crown-group brachiopods, in line with previous interpretations.

Micrina and *Mickwitzia* are the most crownwards of the tommotiids, but beyond this, their position is somewhat difficult to pin down; certain analytical configurations reconstruct them as stem-brachiopods; others place them closer to the discinids, the lingulids or the craniiforms. *Helomedusa* is commonly associated closely with *Mickwitzia*, reflecting the similarities emphasized by Holmer and Popov in Williams et al. (2007), but plots instead within the Craniiforms under certain analytical conditions, in line with earlier interpretations (Williams et al., 2000).

Hyoliths Hyoliths are interpreted as stem-group Brachiopods, which refines the broader phylogenetic position proposed by Moysiuk et al. (2017). This is to say, they sit closer to brachiopods than the phoronids do, but no analysis places them within the Brachiopod crown group.

Hyoliths thus represent derived tommotiids, and are the closest relatives to the Brachiopod crown group.

Supplementary Figures

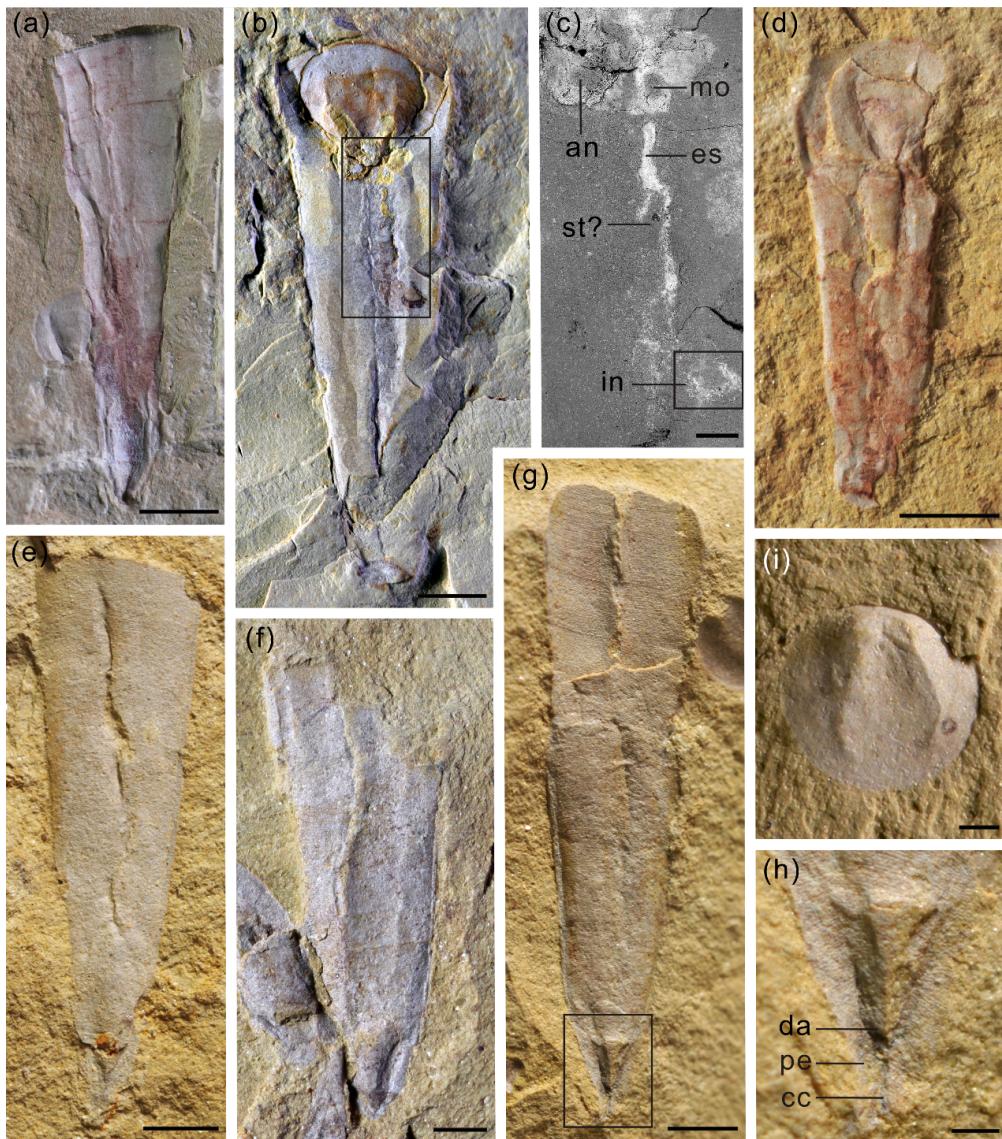


Fig. S1. *Pedunculotheca diania* Sun, Zhao et Zhu gen. et sp. nov. from the Chengjiang Biota, Yunnan Province, China. (a) NIGPAS 166601, external mould of dorsum with dorsal apex and pedicle foramen. (b) NIGPAS 166597, preserving conical shell, operculum and internal soft tissue, showing a compressed elliptic cross-section; backscatter electron micrograph of boxed region shown in (c). (d) NIGPAS 166599b, counterpart, juvenile conical shell with operculum showing two longitudinal ventral grooves and

circular larval shell. (e) NIGPAS 166602, conical shell with incomplete attachment structure. (f) NIGPAS 166598, broken shell with two ventral furrows and incomplete attachment structure. (g) NIGPAS 166596, incomplete shell with one medial ventral furrow and short attachment structure with coelomic cavity; detail of boxed region shown in (h). (i) NIGPAS 166603, exterior of operculum. Scale bars: 2mm (for a, b and e-g); 500 µm (for c, h and i).

Abbreviations: an = anus, cc = coelomic cavity, da = dorsal apex, es = esophagus, in = intestine, mo = mouth, pe = pedicle, st = stomach.

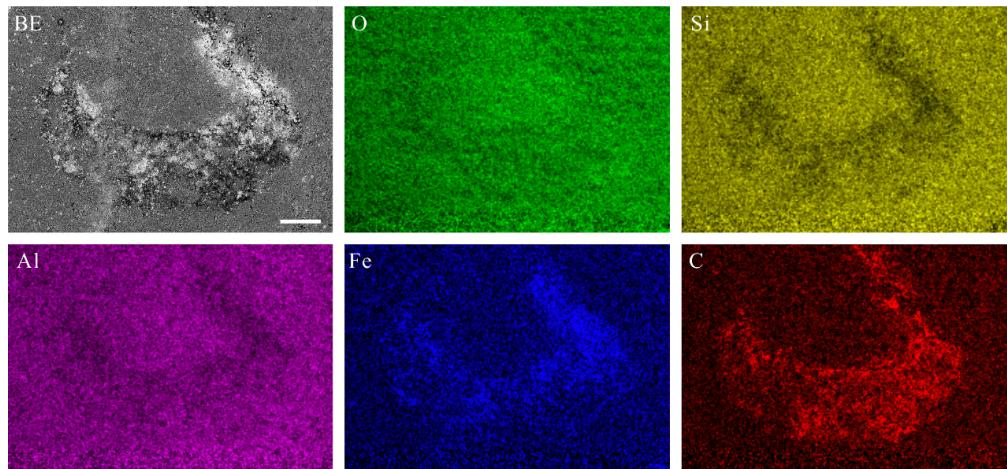


Fig. S2. Elemental distribution in the gut of *Pedunculotheca diania* Sun, Zhao et Zhu gen. et sp. nov. NIGPAS 166597. Region corresponds to boxed region in Fig. S1c. Scale bar = 100 μm .

Abbreviations: BE = backscatter electron image, O = Oxygen, Si = Silicon, Al = Aluminium, Fe = Iron, C = Carbon.

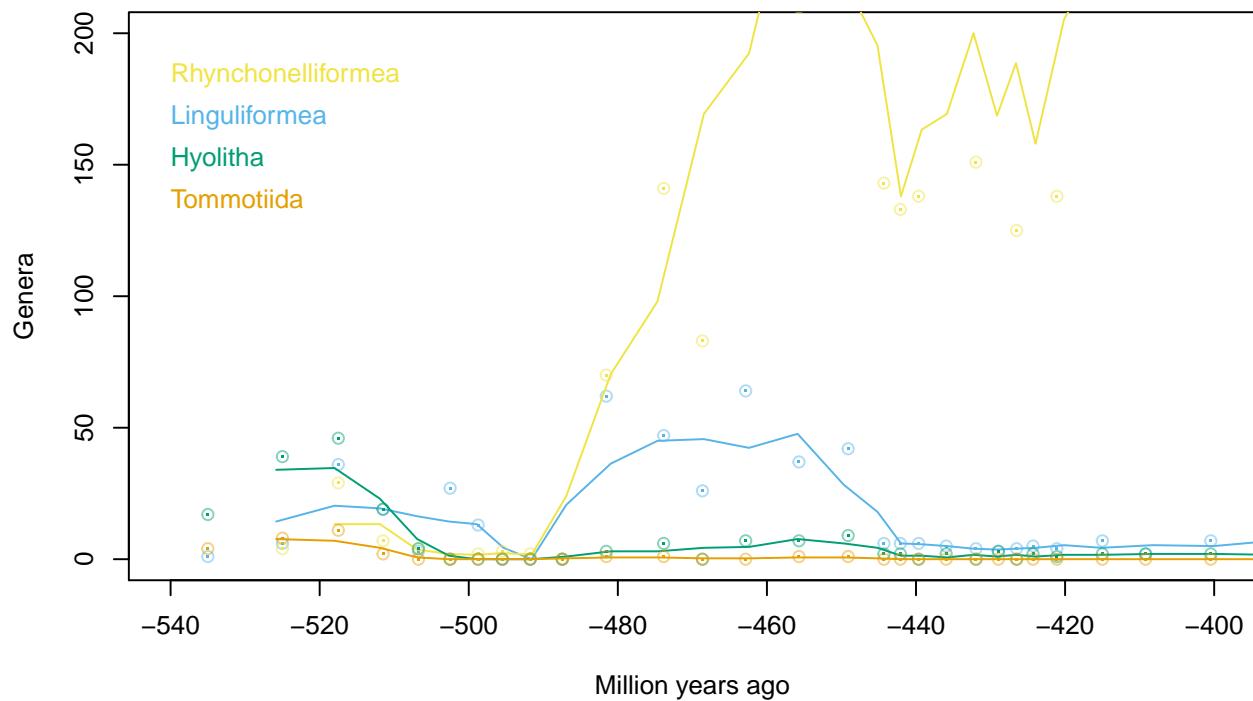


Fig. S3. Global diversity of brachiopods through the Paleozoic. Points represent number of genera reported in each time bin; lines represent rolling mean diversity over three consecutive time bins. Data from Paleobiology database.

Supplementary Table

NIGPAS Specimen numbers	Fossil locality	Coordinates
166593, 166617	Shankou Village, Anning	24°49'53" N, 102°24'47.9" E
166594, 166595	Yaoying Village, Wuding	25°36'01.2" N, 102°20'04.6" E
166596–166616	Ma'anshan Village, Chengjiang	24°40'37.2" N, 102°58'40.2" E

Table S1. Provenance of fossil material. Individuals from the Yaoying section are usually bigger, with a thicker body wall, and have a smaller ratio of apertural width to shell length than specimens from other areas. In the absence of other differentiating features, we consider these deviations to represent ecophenotypical variation within a single species, perhaps reflecting the increased energetics and predation pressure that accompany the shallower water depth reported at the Yaoying section (Zhao et al., 2012).

Bibliography

- Archie, J. W. (1989). Homoplasy excess ratios: new indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the Consistency Index. *Systematic Zoology*, 38(3):253, doi:10.2307/2992286.
- Balthasar, U. (2008). *Mummpikia* gen. nov. and the origin of calcitic-shelled brachiopods. *Palaeontology*, 51(2):263–279, doi:10.1111/j.1475-4983.2008.00754.x.
- Balthasar, U. and Butterfield, N. J. (2009). Early Cambrian soft-shelled brachiopods as possible stem-group phoronids. *Acta Palaeontologica Polonica*, 54(2):307–314, doi:10.4202/app.2008.0042.
- Bassett, M. G., Popov, L. E., and Holmer, L. E. (2001). Functional morphology of articulatory structures and implications for patterns of musculature in Cambrian rhynchonelliform brachiopods. In Brunton, H., Cocks, R. M., and Long, S. L., editors, *Brachiopods, Past and Present*, pages 163–176. Taylor & Francis.
- Brazeau, M. D., Guillerme, T., and Smith, M. R. (2018). An algorithm for morphological phylogenetic analysis with inapplicable data. *Systematic Biology*, doi:10.1101/209775.
- Brazeau, M. D., Smith, M. R., and Guillerme, T. (2017). MorphyLib: a library for phylogenetic analysis of categorical trait data with inapplicability. doi:10.5281/zenodo.815371.
- Budd, G. E. and Jensen, S. (2000). A critical reappraisal of the fossil record of the bilaterian phyla. *Biological Reviews*, 75(2):253–295.
- Carlson, S. J. and Cohen, B. L. (2009). Separating the crown from the stem: defining Brachiopoda and Pan-Brachiopoda delineates stem-brachiopods. *Geological Society of America Annual Meeting Abstract Program*, 41:562.
- Goloboff, P. A. (1997). Self-weighted optimization: tree searches and character state reconstructions under implied transformation costs. *Cladistics*, 13(3):225–245, doi:10.1111/j.1096-0031.1997.tb00317.x.
- Goloboff, P. A. (1999). Analyzing large data sets in reasonable times: solutions for composite optima. *Cladistics*, 15(4):415–428, doi:10.1006/clad.1999.0122.
- Goloboff, P. A. (2014). Extended implied weighting. *Cladistics*, 30(3):260–272, doi:10.1111/cla.12047.
- Goloboff, P. A. and Catalano, S. A. (2016). TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics*, 32(3):221–238, doi:10.1111/cla.12160.
- Holmer, L. E., Pettersson Stolk, P. S., Skovsted, C. B., Balthasar, U., and Popov, L. E. (2009). The enigmatic early Cambrian *Salanygolina* – A stem group of rhynchonelliform chileate brachiopods? *Palaeontology*, 52(1):1–10, doi:10.1111/j.1475-4983.2008.00831.x.
- Holmer, L. E., Popov, L. E., and Bassett, M. G. (2014). Ordovician–Silurian Chileida—first post-Cambrian records of an enigmatic group of Brachiopoda. *Journal of Paleontology*, 88(3):488–496, doi:10.1666/13-104.
- Holmer, L. E., Popov, L. E., Pour, M. G., Claybourn, T., Zhang, Z.-L., Brock, G. A., and Zhang, Z.-F. (2018a). Evolutionary significance of a middle Cambrian (Series 3) *in situ* occurrence of the pedunculate rhynchonelliform brachiopod *Nisusia sulcata*. *Lethaia*, doi:10.1111/let.12254.

- Holmer, L. E., Skovsted, C. B., Brock, G. A., Valentine, J. L., and Paterson, J. R. (2008). The Early Cambrian tommotiid *Micrina*, a sessile bivalved stem group brachiopod. *Biology Letters*, 4:724–728, doi:10.1098/rsbl.2008.0277.
- Holmer, L. E., Zhang, Z.-F., Topper, T. P., Popov, L. E., and Claybourn, T. M. (2018b). The attachment strategies of Cambrian kutoarginate brachiopods: the curious case of two pedicle openings and their phylogenetic significance. *Journal of Paleontology*, 92(1):33–39, doi:10.1017/jpa.2017.76.
- Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, 50(6):913–925, doi:10.1080/106351501753462876.
- Maddison, W. P. (1993). Missing data versus missing characters in phylogenetic analysis. *Systematic Biology*, 42(4):576–581, doi:10.1093/sysbio/42.4.576.
- Moysiuk, J., Smith, M. R., and Caron, J.-B. (2017). Hyoliths are Palaeozoic lophophorates. *Nature*, 541(7637):394–397, doi:10.1038/nature20804.
- Nixon, K. C. (1999). The Parsimony Ratchet, a new method for rapid parsimony analysis. *Cladistics*, 15(4):407–414, doi:10.1111/j.1096-0031.1999.tb00277.x.
- Rannala, B., Zhu, T.-Q., and Yang, Z.-H. (2012). Tail paradox, partial identifiability, and influential priors in Bayesian branch length inference. *Molecular Biology and Evolution*, 29(1):325–335, doi:10.1093/molbev/msr210.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M. A., and Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3):539–42, doi:10.1093/sysbio/sys029.
- Sereno, P. C. (2007). Logical basis for morphological characters in phylogenetics. *Cladistics*, 23(6):565–587, doi:10.1111/j.1096-0031.2007.00161.x.
- Skovsted, C. B., Betts, M. J., Topper, T. P., and Brock, G. A. (2015). The early Cambrian tommotiid genus *Dailyatia* from South Australia. *Memoirs of the Association of Australasian Palaeontologists*, 48(1):1–117.
- Skovsted, C. B., Holmer, L. E., Larsson, C. M., Höglström, A. E. S., Brock, G. A., Topper, T. P., Balthasar, U., Stolk, S. P., and Paterson, J. R. (2009). The scleritome of *Paterimitra*: an Early Cambrian stem group brachiopod from South Australia. *Proceedings of the Royal Society B: Biological Sciences*, 276:1651–1656, doi:10.1098/rspb.2008.1655.
- Smith, M. R. (2017). Quantifying and visualising divergence between pairs of phylogenetic trees: implications for phylogenetic reconstruction. *bioRxiv*, doi:10.1101/227942.
- Smith, M. R. (2018). TreeSearch: phylogenetic tree search using custom optimality criteria.
- Streng, M., Butler, A. D., Peel, J. S., Garwood, R. J., and Caron, J.-B. (2016). A new family of Cambrian rhynchonelliformean brachiopods (Order Naukatida) with an aberrant coral-like morphology. *Palaeontology*, 59(2):269–293, doi:10.1111/pala.12226.
- Sun, H.-J., Smith, M. R., Zeng, H., Zhao, F.-C., Li, G.-X., and Zhu, M.-Y. (2018). Hyoliths with pedicles constrain the origin of the brachiopod body plan. page submitted.
- Vogt, L. (2017). The logical basis for coding ontologically dependent characters. *Cladistics*, doi:10.1111/cla.12209.
- Wiens, J. J. (1998). Does adding characters with missing data increase or decrease phylogenetic accuracy? *Systematic Biology*, 47(4):625–640.
- Wiens, J. J. (2003). Missing data, incomplete taxa, and phylogenetic accuracy. *Systematic Biology*, 52(4):528–538, doi:10.1080/10635150390218330.

- Williams, A., Carlson, S. J., Brunton, C. H. C., Holmer, L. E., Popov, L. E., Mergl, M., Laurie, J. R., Bassett, M. G., Cocks, L. R. M., Rong, J.-Y., Lazarev, S. S., Grant, R. E., Racheboeuf, P. R., Jin, Y.-G., Wardlaw, B. R., Harper, D. A. T., and Wright, A. D. (2000). Linguliformea, Craniiformea, and Rhynchonelliformea (part). In *Treatise on Invertebrate Paleontology, Part H, Brachiopoda (Revised)*, volume 2 & 3, pages 1–919. Geological Society of America & Paleontological Institute.
- Williams, A., Popov, L. E., Holmer, L. E., and Cusack, M. (1998). The diversity and phylogeny of the paterinate brachiopods. *Palaeontology*, 41:221–262.
- Williams, A., Racheboeuf, P. R., Savage, N. M., Lee, D. E., Popov, L. E., Carlson, S. J., Logan, A., Luter, C., Cusack, M., Curry, G. B., Wright, A. D., Harper, D. A. T., Cohen, B. L., Cocks, L. R. M., MacKinnon, D. I., Smirnova, T. N., Baker, P. G., Carter, J. L., Gourvennec, R., Mancenido, M. O., Brunton, C. H. C., Dong-Li, D.-S., Boucot, A. J., Bassett, M. G., Alvarez, F., Holmer, L. E., Mergl, M., Emig, C. C., Rubel, M., and Jia-Yu, J.-R. (2007). Supplement. In *Treatise on Invertebrate Paleontology, Part H, Brachiopoda (Revised)*, volume 6, pages 2321–3226. Geological Society of America & Paleontological Institute.
- Zhang, C., Rannala, B., and Yang, Z.-H. (2012). Robustness of compound Dirichlet priors for Bayesian inference of branch lengths. *Systematic Biology*, 61(5):779–84, doi:10.1093/sysbio/sys030.
- Zhang, Z.-F., Holmer, L. E., Ou, Q., Han, J., and Shu, D.-G. (2011). The exceptionally preserved Early Cambrian stem rhynchonelliform brachiopod *Longtancunella* and its implications. *Lethaia*, 44(4):490–495, doi:10.1111/j.1502-3931.2011.00261.x.
- Zhang, Z.-F., Li, G.-X., Holmer, L. E., Brock, G. A., Balthasar, U., Skovsted, C. B., Fu, D.-J., Zhang, X.-L., Wang, H.-Z., Butler, A. D., Zhang, Z.-L., Cao, C.-Q., Han, J., Liu, J.-N., and Shu, D.-G. (2014). An early Cambrian agglutinated tubular lophophorate with brachiopod characters. *Scientific Reports*, 4:4682, doi:10.1038/srep04682.
- Zhao, F.-C., Hu, S.-X., Caron, J.-B., Zhu, M.-Y., Yin, Z.-J., and Lu, M. (2012). Spatial variation in the diversity and composition of the Lower Cambrian (Series 2, Stage 3) Chengjiang Biota, Southwest China. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 346–347:54–65, doi:10.1016/j.palaeo.2012.05.021.
- Zhao, F.-C., Smith, M. R., Yin, Z.-J., Zeng, H., Li, G.-X., and Zhu, M.-Y. (2017). *Orthozanclus elongata* n. sp. and the significance of sclerite-covered taxa for early trochozoan evolution. *Scientific Reports*, 7(1):16232, doi:10.1038/s41598-017-16304-6.