Earth and Environmental Science Transactions of the Royal Society of Edinburgh, **98**, 311–325, 2008 (for 2007)

Sampling taxa, estimating phylogeny and inferring macroevolution: an example from Devonian terebratulide brachiopods

Sandra J. Carlson[1] and Paul C. Fitzgerald[2]

[1] Department of Geology, University of California, One Shields Avenue, Davis, California, 95616–8605, USA
E-mail: Carlson@geology.ucdavis.edu

[2] Department of Biology, Northern Virginia Community College, 1000 Harry Flood Byrd Highway, Sterling, VA 20164, USA
E-mail: pfitzgerald@nvcc.edu

ABSTRACT: This study explores the implications of different decisions about taxon sampling in studies of phylogeny and macroevolution – how would phylogenetic results differ as decisions about inclusion of taxa differed? The focus is on investigating phylogenetic relationships among families of Devonian terebratulides, and include all 71 named genera in our analyses. Subsets of taxa were experimented with the most complete morphologically from fossil specimens; those that occur earliest in the stratigraphic record; those that include only the name-bearing genera from each family. Including only the most completely known genera produces a result essentially similar to one including all genera, even those for which less than half the characters can be coded. Including only the stratigraphically earliest genera produces a result at odds with the other analyses. Including only name-bearers, representing 13% of all genera, produces a result generally similar to the analysis including all taxa. None of the results of these phylogenetic experiments involving subsets of genera corresponds strongly with the recently revised classification in the *Treatise on Invertebrate Paleontology*, but general similarities can be discerned. The lack of strong correspondence between classification and several different experimental phylogenetic hypotheses could be ascribed to their different overall goals, and highlights the potential dangers of ascribing evolutionary significance to simple counts of taxa, particularly families, as warned by Alwyn Williams 50 years ago.

KEY WORDS: Centronelloidea, character evolution, classification, fossils, morphology, phylogenetic inference, stratigraphic position, Stringocephaloidea

Half a century ago, Alwyn Williams (1957) wrote a provocative paper on evolutionary rates in brachiopods, as a response, of sorts, to G. G. Simpson's *The Major Features of Evolution*, which had been published a few years before. With such a far-reaching title, Simpson's book was enjoying broad appeal among palaeontologists, both invertebrate and vertebrate, and Williams (1957) felt that some of the generalisations about evolution expressed in the book did not accurately reflect what he observed among brachiopods. In particular, Simpson's statement that carnivores may have evolved up to ten times faster than pelecypods prompted Williams to investigate the source and reliability of the supporting data. After comparing the numbers of new brachiopod genera named in 1894, 1929, and 1956 from selected geological time periods, Williams determined that estimates of evolutionary rate can vary significantly as calculated at different points over historical time. As ever-increasing numbers of systematic palaeontologists study and name new taxa from particular time periods (e.g., Ordovician) and particular regions of the world (e.g., United States), research focus in the field changes from decade to decade. These changes can strongly influence the shape of long-term patterns generated from counts of taxa over time, and the evolutionary interpretations such patterns invite. Williams (1957, p. 17) concluded thus: 'for this phylum [Brachiopoda] at least the raw data [on diversity] can change significantly from year to year dependent upon systematic opinion and regional influences . . .' And, following logically, 'there is a real danger that the student will lose sight of the tenuous and arbitrary nature of most of the data used in the compilation of such [diversity] charts, for there is always a tendency to accept numbers as the only worthwhile facts in papers of this kind.' It is noteworthy that Williams was keenly aware, fifty years ago, of the many and varied significant biases inherent in quantifying large-scale patterns in evolution, and sought to warn others of the dangers of such practice, pursued uncritically. Raup (1972; also Raup 1979; Raup & Marshall 1980) later took up this theme and applied it fruitfully to estimates of global Phanerozoic species-level diversity (Valentine 1969). Mistaking noise for signal is a common error, one that is difficult to avoid, particularly when the noise mimics patterns that one expects. Such practice has undesirable consequences when inferring process from pattern in macroevolutionary studies.

If estimates of evolutionary rate change significantly with systematic practice, one can hypothesise that estimates of evolutionary relationships, expressed both as classifications and as phylogenies, may also change significantly. As more fossil taxa are discovered and named, our understanding of morphological diversity encompassed by the clade changes. Yet, decisions must be made about the inclusion or exclusion of taxa of particular rank in any macroevolutionary phylogenetic study involving fossils and morphology because it is

 doi:10.1017/S1755691007078437

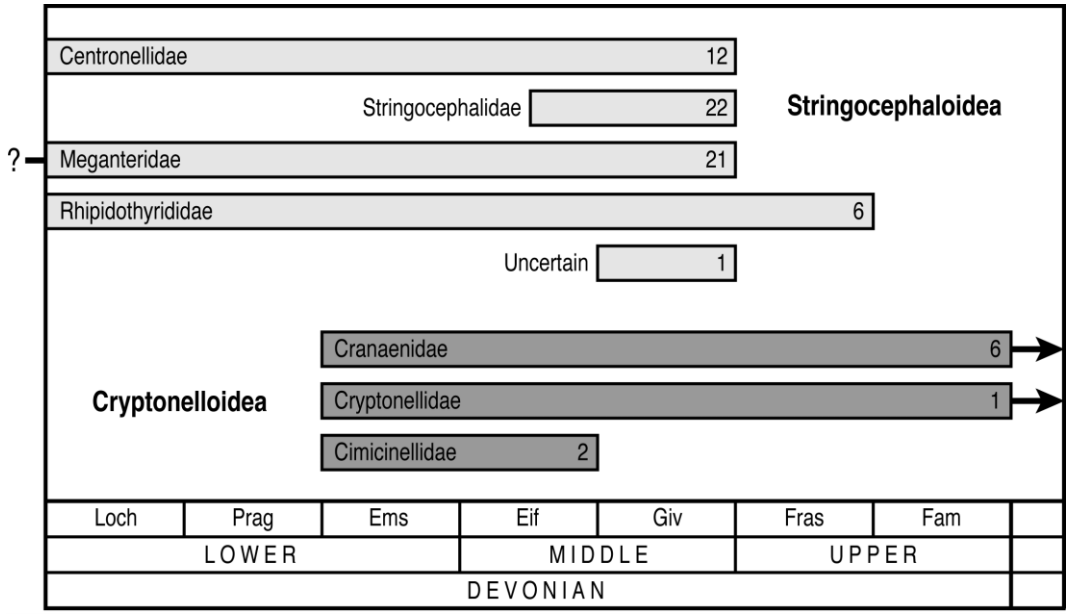

Figure 1 Stratigraphic ranges of the eight Devonian terebratulide families, according to the revised *Treatise on Invertebrate Paleontology* (Lee *et al.* 2006). All Devonian stages are scaled to the same length. Numbers in the range bars indicate the total number of Devonian genera in each family.

not possible to sample truly exhaustively. Does one include all named taxa, knowing that some fossil taxa have not yet been discovered and named? Include all, even though morphological character data may not be sufficient to support robust phylogenetic conclusions? If one chooses not to include all, how are taxa to be sub-sampled? On what criteria are some selected and others not selected?

These questions are considered by investigating patterns of phylogenetic relationships among the earliest (Devonian) members of the most diverse clade of living brachiopods, the Terebratulida. In particular, we experiment with a variety of taxonomic sampling decisions are experimented with (see Hillis 1998), involving inclusion or exclusion of taxa, to determine the effect such decisions have on hypotheses of phylogenetic relationship. Specifically, results obtained from an analysis of all named Devonian terebratulide genera are compared and interpreted with those from analyses including only the most completely known members of each family, the stratigraphically earliest members of each family, and the name-bearers for each family. What differences result? What evolutionary significance do they imply?

1. Materials

Terebratulide brachiopods provide an excellent study group for phylogenetic research. At approximately 100 genera, they are the most diverse extant group. They are also one of the longest-lived groups, with a fossil record over 400 million years long; they experience and rebound from multiple extinction events during this time. Terebratulide phylogeny has been investigated relatively little over the past century (Cloud 1942; Boucot 1959; Boucot *et al.* 1963; Stehli 1965; Boucot & Wilson 1994). Terebratulides are monophyletic (Carlson 1995; Carlson & Leighton 2001), distinguished from their apparent closest relatives, the athyrididines, largely by the possession of a loop-shaped, mineralised, lophophore support. Terebratulides have been claimed to originate by neoteny (Williams 1956; Williams & Wright 1961), an untested hypothesis raising questions about the possible role of heterochrony in their evolution.

The first terebratulides appear in the Lochkovian (Cloud 1942; Lee *et al.* 2006); in order to focus on the origin and early evolution of the clade, the present study was limited to Devonian terebratulides only. Seventy-one genera in eight families and two superfamilies – Stringocephaloidea and Cryptonelloidea – exist in the Devonian, according to Lee *et al.* (2006); the most diverse stringocephaloids are confined to the Devonian (Fig. 1). Genera are referred to according to their familial affiliation in the analyses that follow. Informal familial names end with the '-id' suffix; informal suprafamilial names end with the '-oid' suffix; informal ordinal names end with the '-ide' suffix.

Genera, not species, were used here as the lowest taxonomic rank in this study. Cooper (1970) discussed extensively the nature of morphological variation among brachiopod taxa and concluded that 'the genus is the unit of their evolution and records many small increments of change. It is important because it summarises and groups species into more understandable and meaningful entities' (Cooper 1970, p. 194). This is due in large part to oversplitting at the fossil species level, a taxonomic practice that appears to have continued for the past several decades.

All forty Devonian terebratulide genera listed by Stehli (1965) were considered in this investigation with the exception of two, *Acrothyris* (synonym of *Geranocephalus*) and *Paranaia* (synonym of *Derbyina*), which are no longer valid. The brachiopod *Treatise* volumes have recently been revised (Lee *et al.* 2006), and preliminary lists of terebratulide genera to be included were generously made available for the present study by D. E. Lee prior to publication.

Genera coded using fossil specimens were personally examined (Fitzgerald 2006), and are currently housed, at the National Museum of Natural History in Washington, DC, the Yale Peabody Museum in New Haven, Connecticut, and the Natural History Museum in London, England. When possible, a single species was used to typify the character states for a genus. When multiple species of a genus were available for personal or literature inspection, the most completely preserved, undistorted single specimen was used. Genera not available for examination, and characters that were not preserved in the available specimens but described elsewhere were coded using primary literature sources and available plates. In these cases, the original genus descriptions were used. When

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

the original description did not contain complete skeletal information, missing characters were coded using the most complete description and plates figured in the literature.

2. Methods

2.1. Phylogenetic analyses

Information on variability was compiled from examination of specimens in the above museums, and from literature sources, on 29 binary and 38 multistate characters (Table 1). All non-applicable characters (e.g., dental plate orientation in taxa that lack dental plates) were coded as '-' and missing or unknown character states as '?' (Table 2).

Phylogenetic analyses were performed using PAUP* version 4.0b10 for Macintosh (Swofford 2002), with characters unordered, given equal weight, and then reweighted based on the value of their rescaled consistency index in the initial analysis. This procedure gives proportionally greater weight to characters (Farris 1969) with higher consistency indices in initial analyses, and minimises noise generated by inconsistent characters. Heuristic searches of the matrix were performed with TBR branch-swapping, 100 replicates with random addition of taxa for each analysis.

Because of the large size of the taxon character matrix in the analysis of all Devonian terebratulide genera, the computational time required to obtain bootstrap support estimates (Felsenstein 1985; Sanderson 1989) for this component of the investigation was impractical. Therefore, for this and other analyses in this study for which the calculation of bootstrap support values was computationally impractical, clade support was calculated using the 'fast bootstrap' method as described by Mort *et al*. (2000). In this technique, search effort is substantially reduced by turning off the branch swapping algorithm and substantially increasing the number of replicate data sets (n=10 000). Typically, bootstrap methods that sacrifice search optimality such as the 'fast bootstrap', as well as others (DeBry & Olmstead 2000), provide clade support values that are proportional, but occasionally significantly lower than values obtained when standard bootstrap analyses are performed. For comparison between more conservative 'fast bootstrap' values and values obtained via traditional search methods, 'fast bootstrap' values are presented for all phylogenetic analyses presented in this study.

In this investigation, the cladograms reported were generated by performing heuristic searches, reweighting the characters by the rescaled consistency index, and repeating the analysis. Because the use of reweighted character weights for bootstrap analysis is highly contentious (Liden 1999), the bootstrap values reported using the 'fast bootstrap' are based on unweighted characters (character weights all equal to one). The values presented are only for clades that are present in both the original re-weighted analysis and the fast bootstrap.

Outgroups were chosen from among the Athyridida on the basis of analyses by Carlson & Leighton (2001), which suggested a sister group relationship with the terebratulides. Six genera were selected because of their morphological similarity to early terebratulides and because their stratigraphic ranges bracketed the first appearance of the terebratulides in the Lochkovian. *Nalivkinia*, an atrypide, was also included as an outgroup because Boucot & Wilson (1994) considered it to represent a probable terebratulide ancestor.

2.2. Phylogenetic experiments

Analysis 1: All genera. All 71 genera were analysed initially, using all seven outgroups (Table 2). This analysis provides a starting point for comparison with the four subsampling experiments described below. It has the advantage of being complete, involving all named genera, but the disadvantage of including so many taxa that the set of 67 morphological characters (Table 1) is unlikely to be able to resolve relationships clearly.

Analysis 2: Most completely known genera. None of the 71 taxa could be coded for all 67 characters because of missing or unknown data, or because of the inapplicability of certain characters. A series of analyses was therefore performed, including subsets of taxa (Table 3) to determine what, if any, effect character completeness has on tree topology. The first subset removed the seven most poorly known taxa; the second subset removed 26 additional taxa, leaving only those 38 genera coded for at least 45 (67%) characters (Fig. 2).

Analysis 3: Comparing 'not applicable' and 'unknown' states. PAUP does not treat '?' for missing or unknown information differently from '-' for 'not applicable' characters in analyses, even though they can be coded differently in the matrix (Table 2). Additional collecting may well lead to the discovery of better, more completely preserved specimens, and the character states coded as missing or unknown might then become known. Characters and states that are not applicable will never become known, no matter how much additional collecting occurs. As such, 'unknown' and 'not applicable' are two fundamentally different kinds of characters and states, and should not be treated in the same way in phylogenetic inference (see Novacek 1992; Jenner & Schram 1999; Jenner 2002). Therefore, in a second character/taxon matrix, all '-' states were changed to '9' as recommended by Maddison & Maddison (2000) and treated as a separate state. The problem with this approach is that 'not applicable' is not homologous, and the forced sharing of this state may inappropriately weight the absence of certain characters; this particular experiment set out to test this problem.

Analysis 4: Stratigraphically earliest genera. Ideally, the phylogenetically most basal (presumed to be primitive) members would be chosen as exemplars to represent the family, because the characters they possess are more likely to characterise the evolutionary origin of the group. Less ambiguity is involved in identifying the stratigraphically earliest members of each family, particularly when analysing a matrix with a low character (67) to taxon (71) ratio. Therefore, we selected 28 genera (Table 3) that first appear in the first or second geological stage in which each family is represented: Lochkovian and Pragian for Centronellidae, Meganteridae, and Rhipidothyrididae; Eifelian for Stringocephalidae; Emsian for Cryptonelloidea.

In a second analysis, an attempt was made to test the effect of stratigraphic position on tree topology by including all 71 ingroup genera, and instead designating the ten Lochkovian terebratulide genera as outgroups, in effect using stratigraphic polarity to determine the direction of character state transformation in younger terebratulides.

Analysis 5: Name-bearing genera per family. The nine genera that bear the names of each family represented in the Devonian are most likely to possess characters that typify the family. These characters may serve to distinguish one family from another, but may have evolved and changed at some time(s) after the stratigraphically earliest (and phylogenetically early-branching) members appear, and thus represent derived states. Nevertheless, a few name-bearing genera alone were chosen for this analysis, for comparison with the other analyses.

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

Table 1 Character set

Valve form	
1. Size	Small, 3–10 mm (0); moderate, 10–25 mm (1); large, 25–50 mm (2); exceptionally large, >50 mm (3); exceptionally small, <3 mm (4).
2. Valve outline	Subcircular (0); elongate-oval (1); subquadrate (2);subtriangular (3); subpentagonal (4); subhexagonal (5).
3. Anterior commissure	Rounded (0); straight (1); lobate (2); crenulate (3).
4. Widest part of shell	Medial 1/3 (0); posterior 1/3 (1); anterior 1/3 (2).
5. Adult valve convexity	Ventribiconvex (0); dorsibiconvex (1); DV=VV (2); plano-convex (3); concavo-convex (4).
6. Adult valve convexity strength	Biconvexity weak (0); moderate (1); strong (2); very strong (3).
7. Beak curvature	Suberect (0); incurved (1).
8. Ventral valve beak rostrate	Not rostrate (0); rostrate (1).
Fold and sulcus	
9. Fold and sulcus	Absent [rectimarginate] (0); unisulcate [DV sulcus](1); uniplicate [DV fold](2).
10. Fold and sulcus distribution	Present on entire valve (0); only on anterior 2/3 (1); only on anterior 1/3 (2).
Valve ornament	
11. Valve ornament	Smooth (0); radial ribs (1); plicae (2).
12. If radial ribs:	Costate (0); costellate (1); capillate (2).
13. Valve ornament distribution	Covers entire valve (0); only anterior 2/3 (1); only anterior 1/3 (2); medial 1/3 only (on fold and sulcus) (3).
14. Ornament strength	Weakly developed (0); strongly developed (1).
15. Growth lines	Absent (0); present, weak (1); present, strong (2).
Shell structure	
16. Shell structure	Impunctate (0); endopunctate (1).
Hinge region	
17. Pedicle opening	Absent (0); present as delthyrium or foramen (1).
18. Delthyrium/foramen	Open (0); discrete deltidial plates (1); conjunct deltidial plates (2); symphytium (conjunct, no seam) (2).
19. If foramen present, located:	Hypothyrid (surrounded by deltidial plates, not in contact with beak) (0); submesothyrid (in contact with deltidial plates and beak)(1); mesothyrid (at point of beak) (2); permesothyrid (behind beak) (3); amphithyrid (4); hypothyrid (5).
20. Pedicle collar	Absent (0); present (1).
21. Cardinal margin (posterior margin of dorsal valve)	Broadly angular to straight [140 degrees +] (0); angular [100–140] (1); sharply angular [<100] (2).
Ventral valve interior	
22. Dental plates	Absent (0); present (1).
23. Dental plate orientation	Subparallel (0); convergent (1); divergent (2).
24. Dental plates	Not uniting with median septum (0); uniting with median septum (1); uniting and forming spondylium (2).
25. Median septum	Absent (0); present (1).
26. Median septum height	Low (0); high (1).
27. Diductor muscle scars	Not visible (0); oriented anterior-posterior (1); laterally (2).
28. Diductor muscle scars	Not forming myophragm (0); forming myophragm (1); Incised (2).
Dorsal valve interior	
29. Cardinal process	Absent (0); present (1).
30. Cardinal process size	Short, reduced (0); long, reduced (1); long, exaggerated (2).
31. Cardinal process shape	Unilobed (0); bilobed (1); multilobed (2).
32. Outer hinge plates [between socket ridge and crural base].	Absent (0); present (1).
33. Outer hinge plate length	Short/weak (0); long/strong (1).
34. Outer hinge plates	Thick/solid/sessile (0); thin/as discrete plates (1).
35. Crura	Present (0); absent (1).
36. Crural extensions	Short (0); long (1); loop (2); spiralia (3).
37. Inner hinge plates (=crural plates)	Absent (0); present (1).
38. Inner hinge plate orientation	Parallel (0); convergent and disjunct (1); convergent and V-shaped (sessile) (2); conjunct, forming septalium (3).
39. Septalium/cardinal plate	Supported by median septum (0); unsupported/free (1).
40. Cardinal plate	Flat (0); concave (1).
41. Cardinal plate (consists of outer hinge plate, crural base, and inner hinge plate)	Imperforate (0); perforate (1).

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

Table 1 *Continued.*

42. Cardinal perforation	Small (0); large (1).
43. Dorsal median septum	Absent (0); present (1).
44. Adductor muscle scars	Not visible (0); anterior-posterior (1); lateral (2).
Loop/spiralia	
45. Lophophore	Plectolophe (0); trocholophe (1); schizolophe (2); zygolophe (3); ptycholophe (4); spirolophe (5).
46. Loop lamellar extensions	Short (<1/2 valve length) (0); long (>1/2 valve length) (1).
47. Loop ascending lamellae	Absent (0); present (1).
48. Lamellae curve from crura:	Anteriorly (0); laterally (1); posteriorly (2).
49. Crural processes	Short (0); extended (1); extended and recurved ventrally (2); united to form ring-like loop (3).
50. Loop terminal points	Absent (0); present, small (1); present, elongated (2).
51. Terminal loop phase:	Acuminate/axial (0); deltiform (1); teloform (2).
52. Transverse band between descending lamellae	Absent (0); present (1).
53. Transverse band width	Narrow (0); wide (1).
54. Transverse band flexure	Weak (0); strong (1); unflexed (2).
55. Vertical plate at echmidium	Absent (0); present (1).
56. Vertical plate	Blade-like (0); bifurcate (1).
57. Vertical plate	Not extended posteriorly (0); extended posteriorly (1).
58. Diverging recurved bands from vertical plate	Absent (0); present (1).
59. Spine-like projections from loop	Absent (0); present (1).
60. Spiralia apices oriented:	Laterally (0); medially (1); dorsally (2); ventrally (3).
61. Jugum	Absent (0); present (1).
62. Jugal processes	Absent (0); present (1).
63. Jugal stem	Absent (0); present (1).
64. Accessory jugal lamellae	Absent (0); present (1).
65. Calcified lophophore support	Absent (0); present (1).
66. Anterior jugum	Unfused (0); fused (1).
67. Cardinal plate	Outer hinge plates not fused (0); Anteriorly fused (1); Anteriorly connected by inner hinge plates (2).

3. Results

3.1. Analysis 1: All genera

The agreement of characters among all Devonian terebratulide genera appears to be good, given the largely resolved topology of the strict consensus of 71 trees (Fig. 3; Table 4). Despite the apparent agreement, it is quite difficult to make generalisations about phylogenetic relationships relative to the existing classification. Terebratulides are monophyletic relative to the seven outgroups; none of the named higher taxa, superfamilies or families, are monophyletic. However, some general statements can be made regarding relationships of these taxa. Two major clades emerge. One includes mostly stringocephalid genera emerging from within a small group of cryptonelloid genera, with the cryptonelloids *Cimicinella* and *Cryptonella* in a derived position within; a handful of meganterids and rhipidothyrids, plus *Centronella* and *Hamburgia*, and the meganterid *Pleurothyrella* are at the base of this clade. The other major clade includes all the centronellids (except *Centronella* itself) and two rhipidothyrids as the largest derived clade, with the meganterids *Scaphiocoelia* and *Paulinella* and the stringocephalid *Ectorensselandia* within and *Prorensselaeria* and then *Cimicinoides* at its base; one clade of four meganterids and two centronellids is the sister group to this clade. A clade of five stringocephalids is the sister group to this derived clade, with the cryptonelloid *Costacranaena* at its base. Most meganterids (and *Leioseptathyris* and *Rhipidothyris*) are basal in a paraphyletic comb to this second major clade.

To summarise by family, meganterids are the most dispersed across the topology; most are basal, but at least four genera are in more derived positions in each of the two major clades. Centronellids, except for *Centronella*, occur in one clade, but are interspersed with other non-centronellid genera. Rhipidothyrids are split in two; three genera in one clade, three in the other. Stringocephalids are also split in two groups, but most genera are clustered into one clade. Cryptonelloids are mostly basal to the second major clade, but are spread throughout the topology.

Nalivkinia consistently groups more closely with the athyridides, rather than the terebratulides. It is quite possible that *Nalivkinia* is ancestral to the terebratulides, as Boucot & Wilson (1994) suggest, but it appears to be a more distant ancestor than the Athyridida.

All ten Lochkovian genera, the stratigraphically earliest terebratulides, appear in relatively derived positions (Fig. 3), opposite to what would be predicted on the basis of stratigraphic polarity.

3.2. Analysis 2: Most completely known genera

Morphological character information that is missing or unknown is common among fossil taxa, where lack of adequate preservation, of possibly rare taxa, may prevent certain states from being known. The first experimental sample excluded the seven least completely known taxa (Fig. 2). The results were not sufficiently distinct from the 71-taxon analysis, discussed above, and will not be discussed further.

The second experiment included only the 38 more completely known taxa (Fig. 2). Using all seven outgroups, outgroup monophyly could not be maintained; *Athryis* and *Nalivkinia* consistently separated from the others. With only *Retzia*, *Rhynchospirina*, *Athyrisina*, and *Nucleospira* as outgroups, terebratulide monophyly was recovered (Fig. 4; Table 4). At the base of this cladogram sit the cryptonelloids

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

Table 2 Taxon by character matrix

Adrenia	010001000-2-0101122011000-000--1010211010-1000000000--0--00-----1-?
Afilasma	1100????0-0---011?3121??????0--10102131011000000011000--00-----1-?
Amphigenia	210022100-0---011020111211000--11102101010010100000--0--00-----1-1
Antistrix	14000000100---0110-001000-000--101020-11111000000000--0-000-----1-?
Asiacranaena	210021002?0---11102011200-000--110021?110-010?????????????-----1-?
Athyris	20202200200---101?3011000-100--1010310101001?5--2----------0110111-
Athyrisina	2000220020100100102001100-??0--111030-100-005--2----------0100011-
Barbarothyris	01000110222-2001111011000-000--101020-11100000000000--10000-----1-1
Beachia	110120100-010001121001000-110--11000101010000101001O--10100-----1-1
Bornhardtina	310022110-0---11120000--0-0-0--111020-10110101021101000---1-----1-0
Brachyzyga	003021001020--01122011000-??0--101020-10110?00000???????????----1-0
Centronella	100002110-0---11103000-00-00101100021-11110000000000--10000-----1-0
Chascothyris	20202110200---01112000--0-000--?01020-1-1101010001011010100-----1-0
Cimicinella	11000???0-0---011???21??0???0--10102121110000011021210 21---0-----1-2
Cimicinoides	121021000-0---11120011000-?00--101020-11110?01000000--10101-----1-?
Cloudella	100001110-0-0101110000--0-110--1010?0-1111010???????--?????-----1-1
Cloudothyris	100021000-????11122011000-111111000210100-010?????????????-----1-?
Costacranaena	113021000-102101122011000-100--111020-10110101000011110-000-----1-?
Cranaena	11000?11020---01123121000-0-0--11102131111000000000111100---0-----1-1
Cryptonella	1102???1020---11121021000-0-0--10102131100000011001210 20---0-----1-?
Cydimia	01000100102-0101122011000-000--1010211010-10000000000010000-----1-?
Derbyina	003021000-100101122011000-100--101020-10100101000000--0--00-----1-1
Ectorensselandia	210021100-0---111??011000-001011000202110-0001000000--10100-----1-?
Elmaria	210022100-0---11112??
Erectocephalus	210012110-0---11120000--100-0--1010211----100??????????????????????
Etymothyris	210022100-110011100011100-000--11102101010110?????????????-----1-1
Eurythyris	150000000-010011121001200-101101100010100-0101010010--10100-----1-1
Geranocephalus	100021010-0---01120011200-001211010203000-000?????????????-----1-1
Globithyris	200122110-0-01011?1101000-000--11?0?1310110?????????????????-----1-1
Guangshunia	1?00??010-11???110?001?00-??1101?1????00??1??????????????????????
Hamburgia	110021100-0---01123011000-100--1010203100-000?????????????-----1-?
Hemistringocephalus	341022110-0---01120000--1100121110020-110-0001021101000---1-----1-?
Hessenhausia	310022111-0---11120000--0-0-0--111020-10110101021101000---1-----1-?
Kaplex	200012110-0---11120000--100-0--1010211----100?????????????-----1-?
Kumbella	210001100-0---0112?00100110012111002101l0-000?????????????-----1-?
Leioseptathyris	100001110-0---0112?0?1?10-??1001????0-000-1??????????????????????
Lievinella	210022100-1101111??001200-11101111020-1010010?????????????-----1-1
Lingshanella	10300111101011011220 11000-??0--1010201010-1?01000000--0--00-----1-?
Maclarenella	21003200100---011231?1000-??0--11102031110O?00001110?0--00-----1-?
Meganterella	050021000-0---011?2011000-??1001000210100-010?????????????-----1-1
Meganteris	210020000-0---011?2001000-0-1201000210--0-0001112121020---0-----1-1
Mendathyris	150001100-0-0111122011000-100--100020-1010010???????--?????-----1-1
Micidus	010001000-2-2101122011000-000--1010211110-0000000000--0--00-----1-?
Mutationella	000021000-0-0101112011000-100--100020-10110100000000--10100-----1-0
Nalivkinia	11002310001100101-2011000-0-???10103??----1?5--1----------2110010-
Nanothyris	110021000-0-1101110011000-000--10102101011010000000--10100-----1-1
Neoglobithyris	10000000-1011010101?01000-100--11?02130---100??100?????????-----1-?
Newberria	210022100-0---11112000--0-120--111020-10010101010000--0-000-----1-0
Nucleospira	100021100-0---001?2010--0-??0--10103??1???0?5--2----------011001?-
Omolonia	30002210??0---0112?000--11001211?10203100-000?????????????-----1-?
Oriskania	110020100-0---111?2011200-111121000?10110-000100?0?0--10100-----1-1
Paracrothyris	100021010-0---01120011200-001211010213000-000?????????????-----1-?
Parastringocephalus	30001211100---11120000--100-1111010211----100?????????????-----1-?
Paulinella	200021000-0-0111121001000-110--111020-10100101000010--10100-----1-1
Pleurothyrella	200022100-110101122011000-00100101020-100-000?????????????-----1-1
Podolella	010021100-0--011111011000-?00--10102101011O?00000000--10100-----1-1
Prionothyris	250021000-011-11122001000-111121000210100-0101000010--10100-----1-1
Proboscidina	000021000-0---1111?011000-11111101021?010-01?????????????-----1-?
Prorensselaeria	200121000-0100111--011200-100--10102100011010???????--?????-----1-0
Pseudobornhardtina	310022110-0---11120001000-0-0--111020-10110101021101000---1-----1-?
Reeftonella	100021??0-0---?11??011000-100--10102101010010?????????????-----1-?
Rensselaeria	210122100-1100111??011100-100--11102101010000101000000--10100-----1-1
Rensselaerina	110122100-0-3111111011-00-100--10102101011010001000100101O100----01-1

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

Table 2 *Continued.*

Rensselandioidea	210022100-0---11112000--0-120--111020-00011101010000--0-000-----1-?
Retzia	110001100-100101120101100-0-0--???03131?10015--2----------0101011-
Rhenorensselaeria	210022100-1101111??001200-11101111020-1010110?????????????-----1-1
Rhipidothyris	120021000-100101121001000-000--10102111100110000000--0---0-----1-1
Rhynchospirina	10000210201001011221110000-0-10?1110313000-105--2----------0100011-
Scaphiocoelia	200041100-110001102011000-110--100020-101001?????????????-----1-1
Septothyris	100021000-0---11121001000-000--10102120111100001 0?0??0??00-----1-1
Stringocephalus	300012110-0---11120000--100-1111010211----1001021101000---1-----1-1
Stringodiscus	30000121100---1112000??
Stringomimus	100021010-0---01120011200-001211010203000-000?????????????-----1-?
Sturtella	110220010-0-0101100011000-??0--1010213110-000???????--?????-----1-1
Subrensselandia	210021000-0---011??011000-100--10102101011010102020001100--00-----1-0
Whitfieldella	102022000-0---001?3011100-000--1010303000-2?5--2----------0100011-
Xana	02004110201001011220110000-?00--111020-110-0000000000--0--00-----1-1
Ziangzhounia	100021010-0---01120011200-001211010203000-000?????????????-----1-?

Table 3 Genera included in experiments

	Centronellidae	Stringocephalidae	Meganteridae	Rhipidothyrididae	Cryptonelloidea
Experiment 1: Most complete	*Eurythyris* *Rensselaerina* *Beachia* *Prionothyris* *Oriskania* *Amphigenia* *Rensselaeria* *Nanothyris* *Centronella*	*Chascothyris* *Ectorensselandia* *Subrensselandia* *Hemistringocephalus* *Stringocephalus* *Bornhardtina* *Hessenhausia* *Pseudobornhardtina* *Newberria*	*Cydimia* *Barbarothyris* *Paulinella* *Mutationella* *Antistrix* *Lingshanella* *Xana* *Micidus* *Adrenia* *Rensselandioidea* *Derbyina* *Meganteris*	*Rhipidothyris* *Septothyris* *Podolella*	*Costacranaena* *Cryptonella* *Cimicinoides* *Cranaena* *Maclarenella*
Experiment 2: Stratigraphically earliest	*Nanothyris* *Rensselaerina* *Oriskania* *Rensselaeria* *Eurythyris* *Beachia* *Prionothyris*	*Stringocephalus* *Subrensselandia* *Bornhardtina* *Kaplax*	*Mutationella* *Cloudella* *Mendathyris* *Podolella* *Prorensselaeria* *Brachyzyga* *Sturtella* *Meganteris* Paulinella	*Lievinella* *Globithyris* *Rhenorensselaeria*	*Cranaena* *Asiacranaena* *Cryptonella* *Cimicinella* *Afilasma*
Experiment 3: Name-bearers (+*Guangshunia*)	*Centronella*	*Stringocephalus*	*Meganteris*	*Rhipidothyris*	*Cryptonella* *Cranaena* *Cimicinella* *Afilasma*

Cranaena and *Maclarenella*, then a clade of five stringocephalids, *Centronella*, a clade of most of the meganterids and *Rhipidothyris*, *Septothyris*, and then the most derived clade, which includes clusters of nearly all the centronellids and four stringocephalids as sister groups, intermixed with five meganterids and three cryptonelloids.

The results are very generally similar to those in the 71-taxon analysis. Because so many taxa in the first major (mostly stringocephalid) clade in Figure 3 are incompletely coded, many of these genera were eliminated from this experiment. Three of the four Lochkovian genera that remain are now in a near-basal position in the most derived clade, but are still not where one would predict given their early stratigraphic occurrence.

3.3. Analysis 3: Comparing 'not applicable' and 'unknown' states

The results of analysing the second character/taxon matrix, where '-' states were coded '9' and treated as a separate state (Fig. 5), are a better fit to and more easily interpreted relative to the current classification. Most meganterids are basal to a clade containing most stringocephalids and most centronellids as sister groups. A clade of three cryptonelloids emerges from within the stringocephalid clade, even though the cryptonelloids appear earlier in the fossil record (Fig. 1), and six remaining meganterids are interspersed in the large, more derived clade; rhipidothyrids are split, with *Rhipidothyris* basal and *Septothyris* highly derived.

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

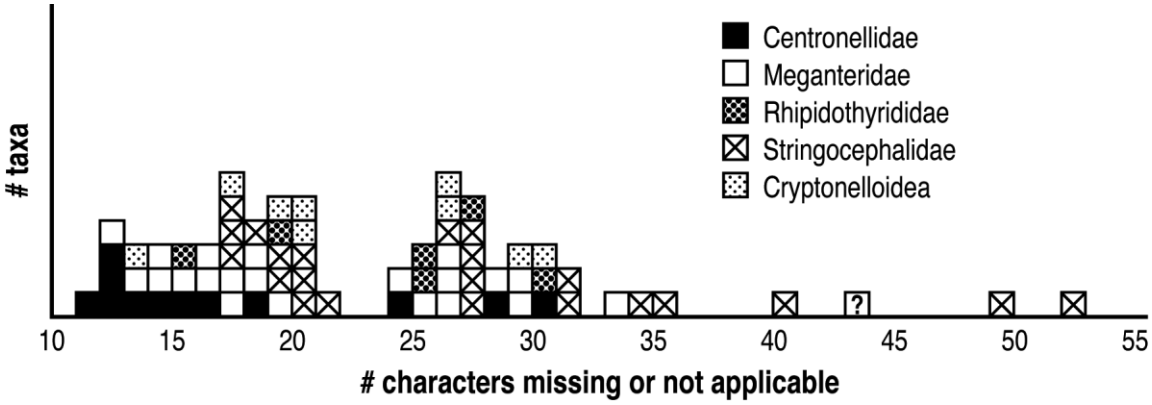

Figure 2 Histogram of the relative completeness of coding of characters.

3.4. Analysis 4: Stratigraphically earliest genera

Including only the earliest occurring genera in an analysis results in a topology rather different from the others discussed so far (Fig. 6; Table 4). Centronellids, with rhipidothyrids, are basal here, instead of the cryptonelloids. Meganterids are then basal to a derived clade of stringocephalids and cryptonelloids/dielasmatoid, with *Globithyris* and three meganterids interspersed. The earliest Lochkovian genera, mostly meganterids (*Mendathyris*, *Podolella*, *Prorensselaeria*, *Brachyzyga*, *Mutationella*, *Cloudella*, *Sturtella*) with two centronellids (*Nanothyris* and *Rensselaerina*) and one rhipidothyrid (*Lievinella*), again branch in the middle of the topology, not at its base, where one might expect.

In the second analysis using stratigraphic polarity, after reweighting characters by their rescaled consistency indices, five of the Lochkovian genera (*Nanothyris*, *Rensselaerina*, *Podolella*, *Mutationella*, and *Prorensselaeria*) remain outside the clade of all other genera, acting as facultative outgroups, but five others nest deeply, in various places in the topology. There appears to be no simple way to force all the earliest taxa to remain basal to the stratigraphically later taxa.

3.5. Analysis 5: Family-name-bearing genera

In the strict consensus of four trees based on the name-bearing genera only (Fig. 7; Table 4), terebratulide monophyly is generally supported, except that a genus of uncertain affinity (*Guangshunia*) groups with the athyridide outgroups, specifically *Rhynchospirina*. *Retzia* appears as the terebratulide sister group, leaving the athyridides as a paraphyletic ancestral group. Centronellids and stringocephalids appear to be sister groups to one another and, as a clade, to a pair of cryptonelloids. In an unresolved polytomy at the base of this clade, lie exemplars from two other cryptonelloid families (cryptonellids and cimicinellids), along with the meganterids and rhipidothyrids. In three of the four trees, rhipidothyrids cluster with the derived well-resolved clade, and the cryptonellids cluster at the base of this clade, with meganterids and cimicinellids unresolved at the base. This result is largely consistent with all the analyses performed thus far, except Analysis 4, involving the stratigraphically earliest genera (Fig. 6).

4. Discussion

4.1. Taxonomic significance

Generalising over all the analyses, stringocephaloids appear to be paraphyletic, with the cryptonelloids emerging from within them. Within the stringocephaloids, meganterids and rhipidothyrids consistently cluster together (with Analysis 4 the only apparent exception), often in a basal position; meganterids are often fragmented. Centronellidae and Stringocephalidae consistently cluster together (except in Analysis 4). Cryptonelloids (cranaenids mostly) are either an imperfect sister group to the stringocephaloids (Fig. 4), or emerge from within or very near to the Stringocephalidae (Figs 3, 5, 6, 7). It is not yet clear, therefore, if the cryptonelloids are the sister group to the stringocephaloids, are a derived subclade within them (as birds are within Dinosauria), or are a paraphyletic group giving rise to them. The data gathered for this study are not sufficiently robust to be able to reject any of these alternatives. These analyses do, however, provide strong support for the decision of Lee *et al*. (2006) to abandon the Centronellidina (see Stehli 1965) as a suborder distinct from the Terebratulidina.

What is the relationship of the phylogenetic hypotheses generated to the revised classification (Lee *et al*. 2006)? Character-based distinctions among many of the higher taxa are quite imprecise (Table 5). Family-level characters are seldom invariant within a family; many exhibit the full range of possible states (e.g., cardinal plates in cryptonelloids). Nevertheless, genera in some subfamilies do appear to cluster consistently together: Eurythyridinae, Kaplexinae, Geranocephalinae, Rensselandiinae, Bornhardtininae, Adreninae, and Rhenorensselaerinae. Others do not: Mutationellinae and Meganteridinae.

The analysis in which 'not applicable' characters were coded as a separate state '9' (Fig. 5) seems to come closest of all to the classification, although the correspondence is not perfect. Weighing the absence of certain characters more than the presence of others seems to reflect at least some of the rationale behind the classification as it currently stands. Identifying this as a potential source of bias, one can begin to ask whether this is a reasonable taxonomic practice or not (Jenner 2002).

Because the morphological character data currently in hand do not resolve phylogenetic relationships with great clarity in any of the several different experiments performed here, and none correspond particularly closely to the current classification, one wonders how morphology can resolve the classification with such apparent clarity. Subfamilies, families, and superfamilies, for example, that could be demonstrated to correspond to definable clades imply different kinds of evolutionary entities than the rather loosely associated collections of genera that can be identified in the cladograms in Figures 3–6. The goals of classification and phylogeny reconstruction are commonly at odds with one another, so it is not unreasonable to expect the results to disagree as well. But, if the quality of the morphological data preclude better resolution in phylogenetic reconstruction, it makes one wonder about their ability to resolve boundaries between higher taxa in a non-arbitrary way, and thus allow macroevolutionary inferences to be made

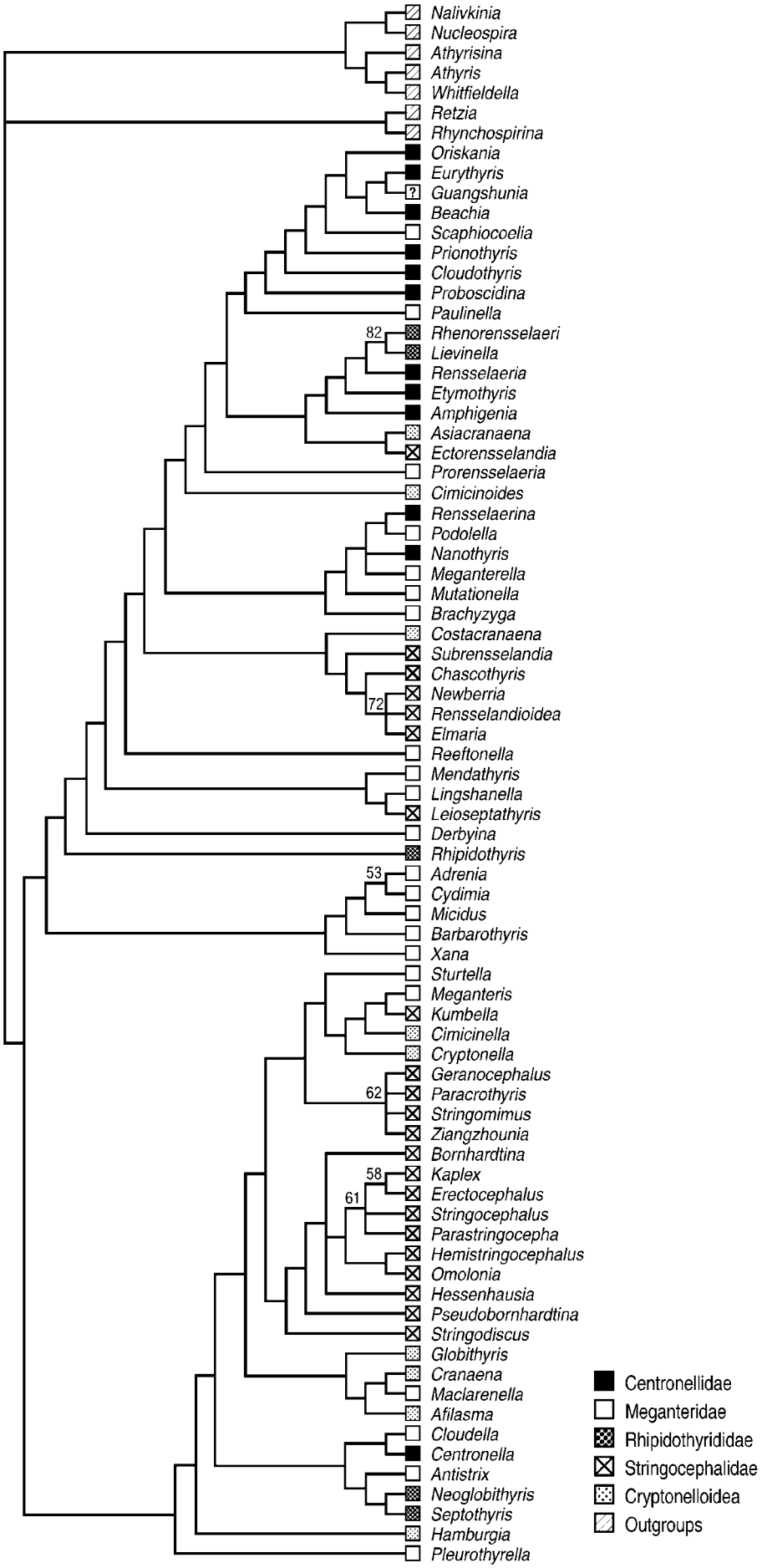

Figure 3 Seventy-one Devonian terebratulide genera as the ingroup, with six athyridide and one atrypide as outgroups. Strict consensus of 71 trees. Symbols indicate the familial affiliations of genera named. Numbers above certain nodes indicate 'fast bootstrap' values supporting those nodes, as in Figures 4, 5, 6, and 7.

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

Table 4 Comparison of Ensemble Consistency Indices (CI) and Ensemble Retention Indices (RI) of analyses. #MP trees refers to the number of equally most parsimonious trees generated per analysis. Experiment 1* utilised a matrix in which all not-applicable characters, designated as '-' in Table 2, were coded as a separate state '9'

	#Genera	#MP trees	#Steps	CI	RI
All	71	71	60·96	0·378	0·734
Experiment 1	38	3	51·70	0·472	0·692
Experiment 1*	38	3	89·88	0·472	0·713
Experiment 2	28	1	63·86	0·522	0·741
Experiment 3	9	4	54·90	0·617	0·729

from those higher taxa (see Sepkoski 1979; Carlson 1996; Alvarez & Modzalevskaya 2001). It is important to try to find out what higher taxa really signify regarding character evolution, a task impossible without at least some kind of phylogenetic analysis for comparison with classification.

4.2. Phylogenetic significance

According to the revision in Lee *et al*. (2006), features of the dental plates, crural plates (inner hinge plates), cardinal plates (the complex of inner and outer hinge plates surrounding the crural base), loop type, and to a lesser extent, median septum and cardinal process are the main characters that can be used to distinguish families of Devonian terebratulides from one another. Table 5 compares characters that generally, but not exclusively, characterise the five major groups compared here. It is clear that some characters are not at all diagnostic for some taxa (nature of the cardinal process in centronellids and stringocephalids), whereas others (nature of the cardinal process in meganterids and rhipidothyrids) can clearly distinguish one group from another. The loop is the character used most often to distinguish superfamilies, but even though most stringocephaloids have acuminate loops, some meganterids have deltiform or teloform loops (which characterise the cryptonelloids and dielasmatoids) and the loop is simply unknown in rhipidothyrids, making it impossible to group them in one superfamily or the other strictly on the basis of this one distinctive character.

What characters structure these cladograms? They vary somewhat from experiment to experiment, but taking just one (Analysis 3) as an example, the characters that support five major nodes in Figure 5 are listed in the figure caption. It is clear that at least some of the characters used to organise the classification play a role in structuring the cladograms (Tables 4, 5, 6), but it is also clear that a number of other morphological characters play significant roles as well (space constraints prevent us from including complete lists of apomorphies defining nodes for the cladogram; lists are available from the authors upon request). No single character or complex of characters reliably distinguishes clades from one another.

When constructing a classification, it is tempting to focus on one or a few characters deemed to be particularly significant, to the exclusion of others, in order to sort differences among taxa in a more or less key-like fashion. Variation in a single character can neatly separate taxa into groups; the appeal of single 'key' characters in establishing classifications is clear. The groups can indicate shared common ancestry if the character is homologous in all, or might inappropriately group together convergent characters if they are merely analogous. In a key to identification, the order of characters examined is important in reaching the correct identification, but the order in the key may be quite different from the order in which these characters actually evolved; a key serves a very different purpose to a phylogenetic hypothesis. One of the advantages of phylogenetic analysis is that all characters are analysed simultaneously, and it becomes possible to investigate the *extent* to which different features support different phylogenetic hypotheses.

How do the samples compare phylogenetically? Different taxonomic samples contain different genera, possessing different combinations of characters. Yet, of all the samples analysed here, only the one involving the stratigraphically earliest genera resulted in a topology different from the others. So the overall results appear to be generally robust to taxon inclusion and exclusion. Specific differences certainly abound, but general patterns emerge consistently. Is one method of sampling to be preferred over another? It depends on the criteria used to evaluate preference. If consistency among the characters supporting nodes is the criterion of choice, it can be evaluated using Ensemble Consistency (CI) and Ensemble Retention (RI) Indices (Table 6), as well as 'fast bootstrap' analyses, to a certain extent. If instead the criterion is consistency with the existing classification, or consistency with relative stratigraphic position, different sampling methods may be employed. Ideally, each of these criteria would point to the same or similar results; if not, an explanation for the source of the disagreement should be sought.

Using name-bearing exemplars alone (Fig. 7) does not appear to be as misleading a sampling method as one might think, given that these taxa are more likely to possess characteristic, derived features of the group. Additional taxa, even if they are poorly known, add some resolution to portions of the tree that cannot be resolved with fewer taxa, as can be seen in Figures 3, 4, 5 and 6.

A clear relationship exists between lower C.I. values and larger numbers of taxa (Sanderson & Donoghue 1989); the Retention Index is not as sensitive to numbers of taxa as is the CI (Table 4). Comparing the CI's (per character) among experiments for the main characters used in the classification (Table 6), some characters that are quite consistent in one experiment are less so in others. However, some characters are likely to be more consistent in all (#31, 51) or less consistent in all (#29, 41).

4.3. Significance of missing data

Are missing data a significant impediment when making inferences about phylogenetic relationships? Apparently not hugely so, as the present results attest. As Kearney & Clark (2003) concluded, missing data lead to a lack of resolution only when characters not missing are in conflict; missing data alone cannot support spurious groupings or lower resolution in the absence of data. Two sets of experiments (#1 [Fig. 4] and #1* [Fig. 5]) attempt to explore the effect of missing data on analyses. While differences are certainly apparent, they are not fundamental; the same general correspondence to the classification is evident, although numerous genera do not 'fit' with other genera in the same family. The coding '9' experiment produced greater correspondence with the classification (although numerous genera still did not 'fit' well), and has a higher Retention Index (Table 5), but the C.I. values were the same as in Analysis 2.

The degree of correspondence desired or required is an important consideration in evaluating the results of these various experiments. If 95% bootstrap support, C.I. values of 0·950, or 95% agreement with stratigraphic FADs and LADs (first or last appearance datum), or 95% correspondence with the existing classification is demanded, then none of these analyses will be seen as having sufficient value. Certain biological systems or phenomena simply do not have statistically

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

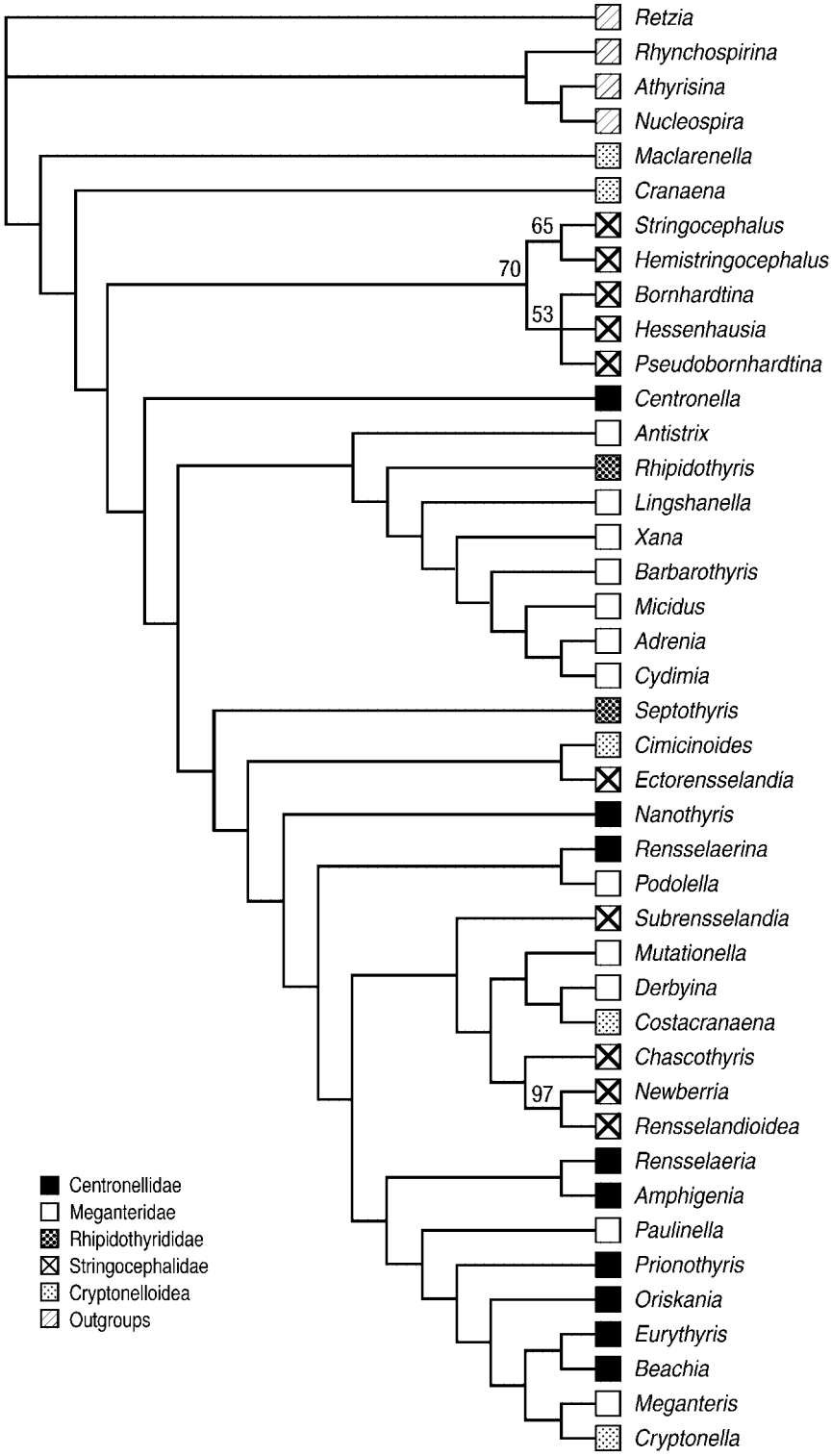

Figure 4 The thirty-eight most complete genera as the ingroup, with four athyridides as outgroups. Strict consensus of three trees.

significant p-values; this does not necessarily invalidate the results of investigations of those systems or phenomena, but requires a different standard of evaluation. The process of the evolution of characters would appear to be one such system; bootstrap values higher than 70–75% are considered to be acceptably 'significant' (Hillis & Bull 1993).

4.4. Stratigraphic significance

If the temporal range of a taxon bears some correspondence to its known stratigraphic range, then stratigraphic position should be able to polarise the direction of evolutionary character change and those taxa that appear earlier in the fossil record would be expected to possess features that are primitive or basal for the clade. However, correspondence between a clade and a higher taxon established prior to phylogenetic analysis is seldom perfect, witness the arguments concerning the definition of Mammalia (e.g., Rowe 1988; Rowe & Gauthier 1992) or Dinosauria (e.g., Fraser *et al.* 2002). Particularly with a non-apomorphy-based (de Queiroz & Gauthier 1990; Doyle & Donoghue 1993) clade definition (either stem-based or crown-based), the earliest members of the clade may or may not possess features that typify the later

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

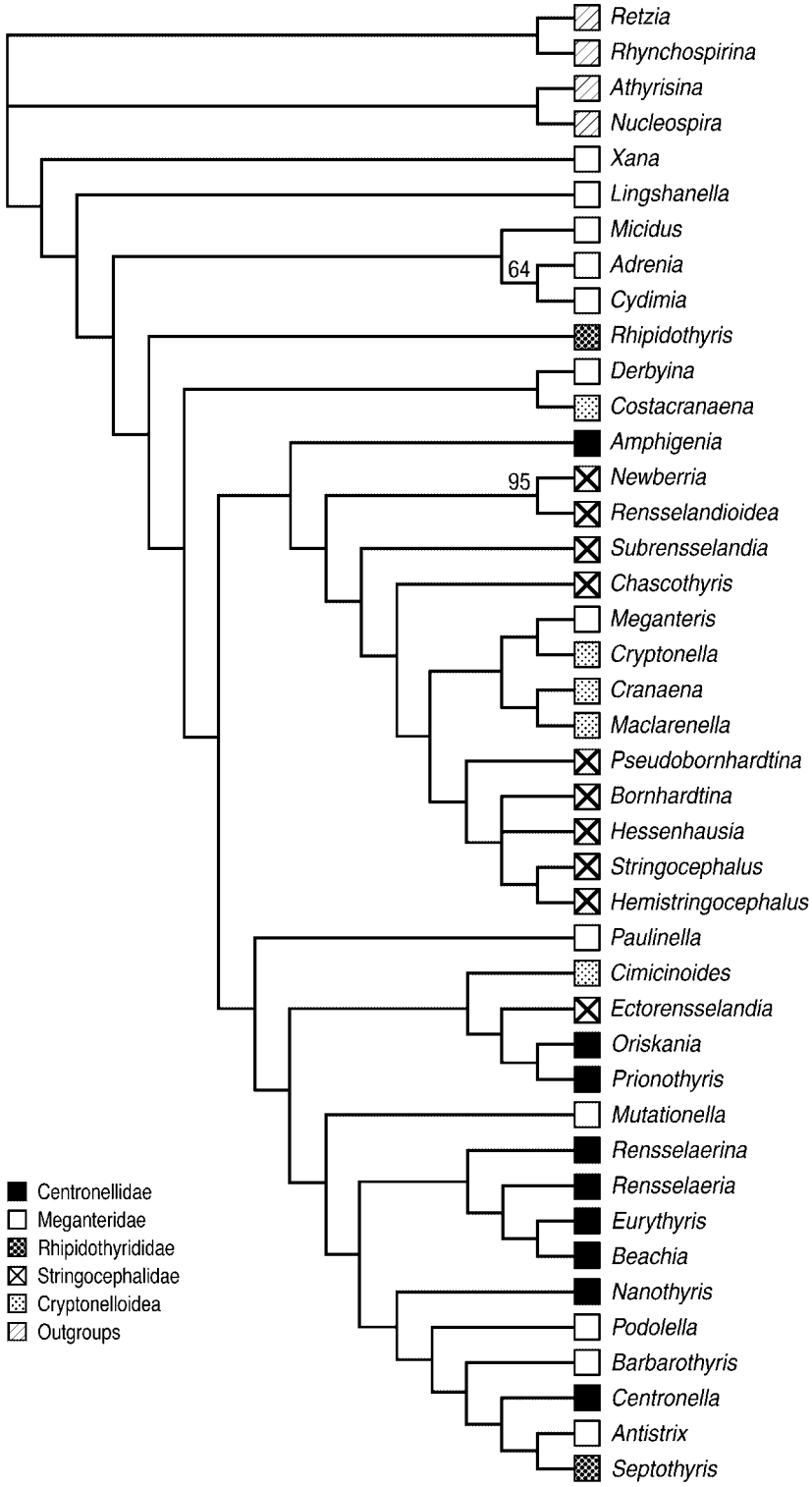

Figure 5 The thirty-eight most complete genera as the ingroup, with four athyridides as outgroups; all 'not applicable' character states coded as a separate state '9' (Analysis 3). Strict consensus of three trees. Twenty characters (2, 36, 45–52, 55, 58–64, 66–67) unite the terebratulides, apart from the outgroups; 16 involve transitions to and from a '9' coded character, the other five concern valve outline, and the nature of the lophophore and lophophore support (loop from spiralia). Three characters (1, 11, 12) unite the two major derived clades and distinguish them from the mostly meganterid taxa at the base of the cladogram; one involves a '9' transition, the other two involve valve size and ornament. Five characters (6, 7, 13, 14, 18) unite the primarily stringocephalid clade; two involve a '9' transition, the others involve the strength of valve convexity and beak curvature, and the nature of the delthyrium and deltidial plates. Four characters (15, 55–57) unite the primarily centronellid clade; two involve a '9' transition, the others involve the presence of growth lines and a vertical plate at the echmidium on the loop.

members; it may be difficult to recognise them morphologically as belonging to a group that contains the later taxa.

Cryptonelloids appear later in time (Emsian compared to Lochkovian) and possess generally different loop morphologies than the stringocephaloids (acuminate vs. deltiform or teloform). They appear mostly in a basal position relative to the stringocephaloids (Figs 3, 4), as a sister group or set of plesions relative to them, but can also appear to emerge

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

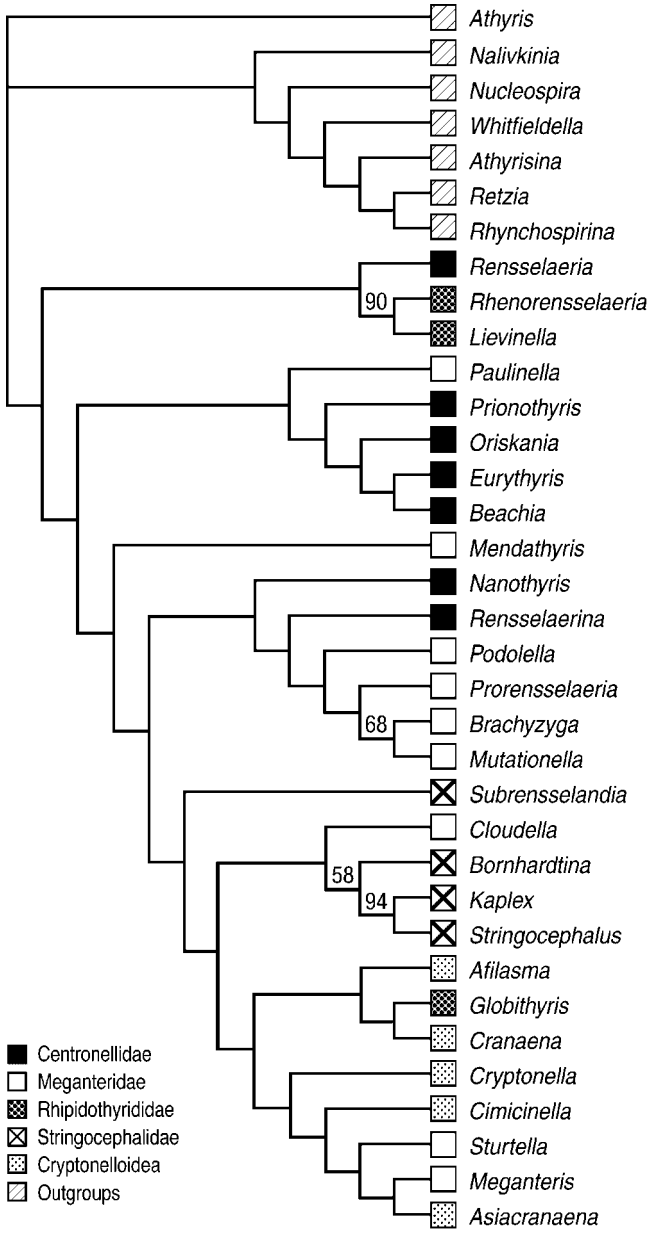

Figure 6 The twenty-eight stratigraphically earliest genera as the ingroup, with six athyridide and one atrypide as outgroups. Single most parsimonious tree.

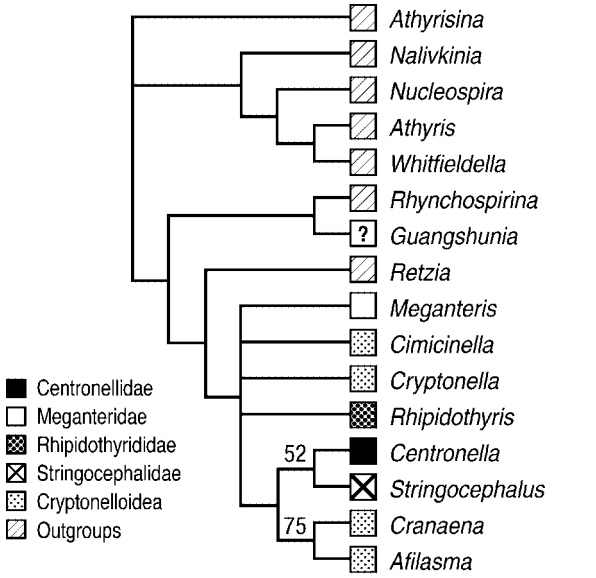

Figure 7 Nine name-bearing genera per family as the ingroup, with six athyridide and one atrypide as outgroups. Strict consensus of four trees.

from within them (Figs 5, 6). This suggests that the characters that distinguish cryptonelloids from stringocephaloids are distinctly different from one another, and depending on the polarity of these characters and the power of the outgroups to polarise them, they can appear either basal or derived. Regardless of which, they are 'pushed away' from one another, in one direction or another, in these topologies.

When analysed with athyridide outgroups, all ten Lochkovian genera, the stratigraphically earliest terebratulides, appear in relatively derived positions (Fig. 3), opposite to what would be predicted on the basis of stratigraphic polarity. This could be caused by any one or all of a number of factors. It is possible that the stratigraphically early taxa might possess, paradoxically, derived characters. This would indicate that stratigraphic polarity is not a good criterion for determining character polarity in evolution. It is possible that the stratigraphically early taxa are less completely known morphologically and that this 'pushes' them to more derived positions in the topology. The present results argue, albeit not strongly, against this possibility, for the reasons discussed in Section 4.3.

It is possible that the athyridides are not the sister group to the terebratulides, making them a poor choice for outgroups to polarise characters, resulting in a topology that is inconsistent with stratigraphic position. It is possible that the stratigraphic ranges of these terebratulides do not yet accurately reflect their evolutionary time range (ghost ranges and lineages; Norell 1992); that further collecting will reveal stratigraphically earlier and phylogenetically more primitive taxa and characters. It is possible that the stratigraphically earliest terebratulides, as adults, possess characters that are more typical of (uncoded here) juvenile athyridides (or atrypides) rather than adult athyridides, as Williams & Wright (1961) and Boucot & Wilson (1994) have suggested. If so, paedomorphosis may have played a significant role in terebratulide evolution, and it will be critically important to evaluate and code juvenile as well as adult character states in all these taxa. It is possible that stratigraphically later terebratulides might have evolved the same character states independently (convergently), and are drawn together in error near the base of the topology, pushing the stratigraphically earliest taxa to more derived positions. Determining the source of incongruence in this particular example of Devonian terebratulides is beyond the scope of the present study, but it will be necessary to resolve this issue eventually, as it is perplexing, counter-intuitive, and not at all uncommon among brachiopods (Carlson 1991).

5. Summary and conclusions

The main purpose in this study was to explore the implications of different decisions about taxon sampling – how would phylogenetic results differ as decisions about inclusion of taxa differed? In other words, how stable are the results to sampling perturbations? To accomplish this goal, results were compared from various taxonomic sampling strategies involving all Devonian terebratulide genera, only the most completely known members of each family, the stratigraphically earliest members of each family, and the name-bearing genus for each family. Missing or not applicable data, resulting in taxa that are coded for fewer than 100% of the characters, do not seem to have a fundamentally important effect on structuring the topologies generated here, or on reducing their overall

https://doi.org/10.1017/S1755691007078437 Published online by Cambridge University Press

Table 5 Summary of a few character complexes relevant to the classification, according to Lee *et al*. 2006. Numbers under the character names refer to character numbers in the matrix and in Table 4

	Centronellidae	Stringocephalidae	Meganteridae	Rhipidothyrididae	Cryptonelloidea
Dental plates (22–24)	Mostly 'obsolete' Spondylium in some	Mostly 'obsolete' Spondylium in some	Present, short	'Obsolete'	Present, short
Crural plates (37–38)	Present, long	Present or absent	Mostly absent	Present	Present, divided or undivided
Cardinal plates (39–42)	Undivided, perforate or discrete	Discrete	Present, discrete	Discrete or fused Septalium in some	Divided or undivided; Perforate or imperforate
Loop (51)	Acuminate	Acuminate, long	Mostly acuminate Some deltiform or teloform	Unknown	Deltiform or teloform
Median septum (25, 39, 43)	Absent (D or V)	Present or absent	Mostly absent	Present in DV	Absent
Cardinal process (29–31)	Present or absent	Present or absent	Present	Absent	Mostly present

Table 6 Comparison of consistency indices per character, as numbered in the character list in Table 1, for some of the main characters used in the revised classification. Boldfaced numbers designate those higher than the Ensemble Consistency Index for the entire tree

Character	All	Exp. 1	Exp. 1*	Exp. 2	Exp. 3
22	**0·400**	0·333	0·333	**0·667**	0·500
23	0·200	0·400	0·250	0·250	0·333
24	1·000aut	1·000aut	0·400	—	—
25	0·333	**0·500**	**0·500**	**1·000**	1·000aut
29	0·071	0·143	0·167	0·167	0·333
30	**0·667**	**0·500**	0·375	**1·000**	0·667aut
31	**0·400**	**1·000**	**0·500**	**0·667**	**1·000**
37	0·053	0·100	0·083	0·125	0·500
38	0·375	**0·750**	0·333	**0·600**	0·600
39	0·100	0·200	0·286	0·333aut	0·500
40	0·091	0·200	0·222	0·250	0·500
41	0·067	0·125	0·250	0·167	0·250
42	0·200	0·250	0·143	0·500	**1·000**
43	0·167	0·143	0·167	0·400	0·400
51	**0·500**	**0·500**	0·429	**0·667**	**1·000**

resolution. None of the cladograms constructed agree particularly well with the classification recently adopted, quite possibly because classification and phylogeny often have different goals and can produce different results. Relative stratigraphic position is not in especially good agreement with cladogram topology.

The morphological data were unfortunately not sufficiently numerous or robust to be able to produce very strongly supported results in any of the taxon sampling experiments. If these data are not sufficient to produce robust results in phylogenetic reconstruction, to what extent can they be trusted as a sound basis for classification? In other words, how reliable is the morphological basis for the names and hierarchy of names in a classification when it is used to make macroevolutionary inferences? Do the higher taxonomic names correspond to clades? Are they based on single or multiple characters? Are they weighted, and if so, how, and on what basis? Are those characters plesiomorphic or homoplastic? There is no way to know until at least some attempt at phylogenetic analysis has been made, which is what this study has attempted to do. Seeking some knowledge about the data in hand, even if they turn out to be relatively weak, is preferable to avoiding gathering knowledge about the data at all.

When counting numbers of taxa to tally up diversity over time, in making inferences about macroevolutionary pattern, as is commonly done in much paleontological literature, no tests for 'robustness' of taxon names exist. Early warnings from Williams and Raup and others of possible biases involved in taxon counting in macroevolutionary investigations, and the use of classifications for purposes other than what they were established to accomplish, appear to have gone largely unheeded until, perhaps, relatively recently (for example, see Peters & Foote 2001). The desire to generalise evolutionary patterns on the broadest possible scale, across space (global), time (Phanerozoic), and taxa (metazoans), is strong and has dominated much paleobiological inquiry for the past several decades (e.g., Sepkoski 1979; Raup & Sepkoski 1982; Jablonski 1986, 1994; Allison & Briggs 1993; Alroy *et al*. 2001; Madin *et al*. 2006). However, the specific implications of different decisions made regarding taxonomic sampling, and the validity of evolutionary conclusions drawn from them, remain poorly known for most clades. Indeed, Williams (1957, p. 17) states that 'even if such data [on taxonomic diversity] are accepted as reasonable reflections of brachiopod evolution, they lead to conclusions different from those published for other phyla,' suggesting that such broad generalisations may have limited value in helping to understand and appreciate the significance of the variety of evolutionary patterns from clade to clade.

To reach a more complete understanding of the evolutionary significance of classifications, it is necessary to evaluate taxa relative to their behavior in phylogenetic analyses, even in analyses that are not exceptionally well resolved. Simply counting taxa in a macrovolutionary study is not sufficient. As Alwyn Williams argued 50 years ago, it is necessary to discover what higher taxa actually signify regarding evolution.

6. Acknowledgments

We thank Maggie Cusack and David Harper for inviting us to contribute to this volume in honour of the many, far-reaching contributions of the late Alwyn Williams. His contributions to the field of brachiopod paleontology are enormously valuable. We thank T. White, C. MacClintock, and S. Butts of the Yale Peabody Museum of Natural History; J. Thompson & M. Florence of the National Museum of Natural History, Washington, D.C.; and S. L. Long of the Natural History Museum, London, who were extremely helpful to PCF while gathering the data for this study. The artistic skills of J. C. Fong are much-appreciated. We gratefully acknowledge the support of the National Science Foundation (EAR 0229897).

7. References

Allison, P. A. & Briggs, D. E. G. 1993. Paleolatitudinal sampling bias, Phanerozoic species-diversity, and the end-Permian extinction. *Geology* **21**, 65–8.

Alroy, J., Marshall, C. R., Bambach, R. K., Bezusko, K., Foote, M., Fursich, F. T., Hansen, T. A., Holland, S. M., Ivany, L. C., Jablonski, D., Jacobs, D. K., Jones, D. C., Kosnik, M. A., Lidgard, S., Low, S., Miller, A. I., Novack-Gottshall, P. M., Olszewski, T. D., Patzkowsky, M. E., Raup, D. M., Roy, K., Sepkoski, J. J., Sommers, M. G., Wagner, P. J. & Webber, A. 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences of the USA* **98**(1), 6261–6.

Alvarez, F. & Modzalevskaya, T. L. 2001. Trends in athyridide diversity dynamics. *In* Brunton, C. H. C., Cocks, L. R. M. & Long, S. L. (eds) *Brachiopods Past and Present. The Systematics Association Special Volume Series* **63**, 212–23. London: Taylor and Francis.

Boucot, A. J. 1959. Brachiopods of the Lower Devonian rocks at Highland Mills, New York. *Journal of Paleontology* **33**, 727–69.

Boucot, A. J., Carter, K. E. & Ives, D. 1963. Relationships of a new lower Devonian terebratuloid (Brachiopoda) from Antarctica. *Bulletins of American Paleontology* **46**, 81–123.

Boucot, A. J. & Wilson, R. A. 1994. Origin and early radiation of terebratuloid brachiopods; thoughts provoked by *Prorensselaeria* and *Nanothyris*. *Journal of Paleontology* **68**, 1002–25.

Carlson, S. J. 1991. A phylogenetic perspective on articulate brachiopod diversity and the Permo–Triassic extinctions. *In* Dudley, E. C. (ed.) *The Unity of Evolutionary Biology*, Vol. **1**, 119–42. Portland, Oregon: Dioscorides Press.

Carlson. S. J. 1995. Phylogenetic relationships among extant brachiopods. Cladistics **11**, 131–97.

Carlson, S. J. 1996. Revision and review of the Order Pentamerida. *In* Copper, P & Jisuo Jin (eds) *Brachiopods: Proceedings of the Third International Brachiopod Congress, 2–5 September 1995*, 53–8. Rotterdam: A.A. Balkema.

Carlson, S. J. & Leighton, L. R. 2001. The phylogeny and classification of the Rhynchonelliformea. *In* Carlson, S. J. & Sandy, M. R. (eds) *Brachiopods Ancient and Modern*, 248–58. New Haven, Connecticut: The Paleontological Society and Yale University.

Cloud, P. E. 1942. Terebratuloid brachiopods of the Silurian and Devonian. *Geological Society of America Special Paper* **38**, 1–182.

Cooper, G. A. 1970. Generic characters of brachiopods. *In* Yochelson, E. L. (ed.) *Proceedings of the North American Paleontological Convention, Field Museum of Natural History, Chicago, September 5–7 1969* Vol. **1**, Part C, 194–263. Lawrence, Kansas: Allen Press.

DeBry, R. W. & Olmstead, R. G. 2000. A simulation study of reduced tree-search effort in bootstrap resampling analysis. *Systematic Biology* **49**, 171–9.

de Queiroz, K. & Gauthier, J. 1990. Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. *Systematic Zoology* **394**, 307–22.

Doyle, J. A. & Donoghue, M. J. 1993. Phylogenies and angiosperm diversification. *Paleobiology* **19**, 141–67.

Farris, J. S. 1969. A successive approximations approach to character weighting. *Systematic Zoology* **18**, 374–85.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–91.

Fitzgerald, P. C. 2006. *Latitudinal diversity gradients, morphological variability, and phylogenetic relationships in Paleozoic terebratulide (Brachiopoda) genera*. Ph.D. dissertation, Department of Geology, University of California, Davis, 237 pp.

Fraser, N. C., Padian, K., Walkden, G. M. & Davis, A. L. M. 2002. Basal dinosauriform remains from Britain and the diagnosis of the Dinosauria. *Palaeontology* **45**, 79–95.

Hillis, D. M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Systematic Biology* **47**, 3–8.

Hillis, D. M. & Bull, J. J. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* **42**, 182–92.

Jablonski, D. 1986. Background and mass extinctions: the alternation of macroevolutionary regimes. *Science* **231**, 129–33.

Jablonski, D. 1994. Extinctions in the fossil record. *Philosophical Transactions of the Royal Society of London, Series B* **344**, 11–17.

Jenner, R. A. 2002. Boolean logic and character state identity: pitfalls of character coding in metazoan cladistics. *Contributions to Zoology* **71**, 67–91.

Jenner, R. A. & Schram, F. R. 1999. The grand game of metazoan phylogeny: rules and strategies. *Biological Reviews* **74**, 121–42.

Kearney, M. & Clark, J. M. 2003. Problems due to missing data in phylogenetic analyses including fossils: A critical review. *Journal of Vertebrate Paleontology* **23**, 263–74.

Lee, D. E., MacKinnon, D. I., Smirnova, T. N., Baker, P. G., Jin Yu-gan & Sun Dong-li. 2006. Order Terebratulida. *In* Kaesler, R. L. (ed.) *Treatise on Invertebrate Paleontology, Part H, Brachiopoda (Revised)*, **Vol. 5**, 1965–2244. Boulder, Colorado and Lawrence, Kansas: Geological Society of America and University of Kansas Press.

Liden, M. 1999. Boots-traps. *Taxon* **48**, 349–50.

Maddison, D. R. & Maddison, W. P. 2000. *MacClade 4, version 4.02, manual*. Sunderland, Massachusetts: Sinauer Associates, Inc.

Madin, J. S., Alroy, J., Aberhan, M., Fursich, F. T., Kiessling, W., Kosnik, M. A. & Wagner, P. J. 2006. Statistical independence of escalatory ecological trends in Phanerozoic marine invertebrates. *Science* **312** (5775), 897–900.

Mort, M. E., Soltis, P. S., Soltis, D. E., Mabry, M. L. 2000. Comparison of three methods for estimating internal support on phylogenetic trees. *Systematic Biology* **49**, 160–71.

Norell, M. A. 1992. Taxic origin and temporal diversity: the effect of phylogeny. *In* Miyamoto, M. M. & Wheeler, Q. D. (eds) *Extinction and Phylogeny*, 89–118. New York: Columbia University Press.

Novacek, M. J. 1992. Fossils, topologies, missing data, and the higher level phylogeny of eutherian mammals. *Systematic Biology* **41**, 58–73.

Peters, S. E. & Foote, M. 2001. Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology* **27**, 583–601.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. *Science* **177**, 1065–71.

Raup. D. M. 1979. Biases in the fossil record of species and genera. *Carnegie Museum of Natural History, Bulletin* **13**, 85–91.

Raup, D. M. & Marshall, L. G. 1980. Variation between groups in evolutionary rates: a statistical test of significance. *Paleobiology* **6**, 9–23.

Raup, D. M. & Sepkoski, J. J. 1982. Mass extinctions in the marine fossil record. *Science* **215**, 1501–3.

Rowe, T. 1988. Definition, diagnosis, and origin of Mammalia. *Journal of Vertebrate Paleontology* **8**, 241–64.

Rowe, T. & Gauthier, J. 1992. Ancestry, paleontology, and definition of the name Mammalia. *Systematic Biology* **41**, 372–8.

Sanderson, M. J. 1989. Confidence limits on phylogenies: the bootstrap revisited. *Cladistics* **5**, 113–29.

Sanderson, M. J. & Donoghue, M. J. 1989. Patterns of variation in levels of homoplasy. *Evolution* **43**, 1781–95.

Sepkoski, J. J. 1979. A kinetic model of Phanerozoic taxonomic diversity: 2, Early Phanerozoic families and multiple equilibria. *Paleobiology* **5**, 222–51.

Simpson, G. G. 1953. *The Major Features of Evolution*. New York: Columbia University Press.

Stehli, F. G. 1965. Paleozoic Terebratulidina. *In* Moore R. C. (ed.) *Treatise on Invertebrate Paleontology*, Part. H, Brachiopoda, Vol. **2**, H730–H762. Boulder, Colorado and Lawrence, Kansas: Geological Society of America and University of Kansas Press.

Swofford, D. L. 2002. PAUP*. *Phylogenetic Analysis Using Parsimony* (**and Other Methods*). Version 4. Sunderland, Massachusetts: Sinauer Associates.

Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. *Palaeontology* **12**, 684–709.

Williams, A. 1956. The calcareous shell of the Brachiopoda and its importance to their classification. *Biological Reviews* **31**, 243–87.

Williams, A. 1957. Evolutionary rates of brachiopods. *Geological Magazine* **94** (3), 201–11.

Williams, A. & Wright, A. D. 1961. The origin of the loop in articulate brachiopods. *Palaeontology* **4**, 149–76.

MS received 8 July 2006. Accepted for publication 4 November 2007.