Assessing conflict between early neornithischian tree topologies

Brown, Emily; Butler, Richard; Barrett, Paul; Maidment, Susannah

DOI:
10.1080/14772019.2022.2032433

Assessing conflict between early neornithischian tree topologies

Emily E. Brown[a,b,*], Richard J. Butler[b], Paul M. Barrett[a] and Susannah C. R. Maidment[a,b]

[a]*Department of Earth Sciences, Natural History Museum, Cromwell Road, London, SW7 5BD, UK;* [b]*School of Geography, Earth & Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK*

SUGGESTED RH—Conflict in neornithischian tree topologies

*Corresponding author. Email: emily.brown2@nhm.ac.uk

Abstract

The phylogenetic relationships of the species commonly referred to as ‘hypsilophodontids’ remains one of the key questions in ornithischian dinosaur research, having profound implications for understanding the origin, evolution and taxonomic compositions of several more recently evolved neornithischian clades. Recent phylogenetic analyses have recovered two conflicting placements for these taxa: (1) primarily within Cerapoda (Ornithopoda + Marginocephalia), as a paraphyletic assemblage of early ornithopods; and (2) primarily outside of Cerapoda, within the clade Thescelosauridae. Here we assess three recent independent neornithischian phylogenetic studies that have recovered topologies congruent with one of these placements. We compare the compositions of these data matrices and test how each of them responds to manipulation of taxa and characters. The positions in which controversial clades are recovered is shown to be highly dependent on the sample of taxa analysed; however, taxon incompleteness or instability is not a contributing factor in altering topology. Character completeness and homoplasy is shown not to significantly alter tree topology either, although these factors can affect resolution. In one matrix investigated, femoral and dental characters are found to provide disproportionate support for the placement of key taxa outside of Cerapoda, and the exclusion of a small number of these characters results in ‘hypsilophodontids’ falling within Ornithopoda. In contrast, matrices that originally recovered ‘hypsilophodontids’ within Cerapoda are comparably more stable, with this array of taxa remaining in a consistent position throughout all analyses. There is still much work to be done to resolve these relationships, but our study provides several suggestions for future analyses with the aim of resolving areas of conflict within the neornithischian tree.

Keywords: Mesozoic; Ornithischia; Neornithischia; phylogenetic relationships; Cerapoda

Introduction

Neornithischians were a diverse and highly successful clade of ornithischian dinosaurs with a fossil record extending from the Middle Jurassic to the end of the Maastrichtian (Sereno 1999; Butler *et al.* 2008; Boyd 2015; Baron *et al*. 2017a; Dieudonné *et al.* 2020). Throughout their evolutionary history, neornithischians gained a worldwide distribution and gave rise to three of the five main ornithischian clades: Ceratopsia, Pachycephalosauria and Iguanodontia (Norman 1984; Cooper 1985; Maryańska & Osmólska 1985; Sereno 1999; Butler *et al.* 2008; Boyd 2015). Although the relationships of these three clades to each other are largely accepted (with ceratopsians and pachycephalosaurs usually united in Marginocephalia, which in turn is the sister clade of Ornithopoda), the relationships of neornithischian taxa that fall outside, or close to the roots, of these main groups are poorly resolved. These taxa tend to be bipedal, small to medium-bodied, and lack the distinctive and often ornate morphological features (e.g. horns, crests, frills) seen in later branching cerapodan clades (Norman *et al.* 2004).

Traditionally, taxa that shared this bauplan and were of uncertain phylogenetic affinities were often assigned to 'Hypsilophodontidae' (e.g. Thulborn 1971; Galton 1973), which became a diverse taxonomic 'waste-basket' with wide temporal and spatial distributions. Although modern cladistic analyses generally agree that this group is para- or polyphyletic (e.g. Scheetz 1999; Sereno 1999; Buchholz 2002; Butler *et al*. 2008; Boyd 2015), this leaves questions as to the phylogenetic placements of these 'hypsilophodontids' (as we will refer to them here) within Neornithischia.

The last decade saw several large-scale independent phylogenetic studies focusing on neornithischian relationships (e.g. Butler *et al.* 2011; Boyd 2015; Dieudonné *et al.* 2016, 2020; Han *et al.* 2018). However, these analyses do not agree on the phylogenetic positions

of the ‘hypsilophodontids’, generally recovering these taxa in two strongly contrasting positions, either diverging before (e.g. Boyd 2015; Madzia *et al.* 2018; Herne *et al.* 2019) or after (e.g. Sereno 1999; Buchholz 2002; Butler *et al.* 2005, 2008; Dieudonné *et al.* 2016, 2020; Han *et al.* 2018) the origins of Marginocephalia and Ornithopoda in the Middle–Late Jurassic. This lack of resolution has impeded our understanding of important character changes throughout the evolutionary history of neornithischians and, consequently, our interpretation of some of the major transitions seen in more deeply-nested cerapodan groups, such as the acquisitions of obligate high-fibre herbivory (Barrett 2014; Mallon & Anderson 2015) and quadrupedality (Maidment & Barrett 2012).

[Insert FIGURE 1 near here]

Here, we assess the two main current hypotheses (Fig. 1) concerning the phylogenetic position of the ‘hypsilophodontids’: hypothesis 1 (H1), that these taxa primarily fall within Cerapoda (Marginocephalia + Ornithopoda) as early-diverging ornithopods; and hypothesis 2 (H2), that the majority are non-cerapodan neornithischians, within the clade Thescelosauridae (Orodrominae + Thescelosaurinae). We investigate how character and taxon sampling has influenced the recovery of these two contrasting hypotheses by comparing and contrasting three recent datasets compiled to examine neornithischian phylogeny, with the aim of aiding future assessments of early neornithischian relationships.

Previous analyses

Most large-scale studies looking at ornithischian and neornithischian relationships over the past two decades have recovered ‘hypsilophodontids’ in a position congruent with H1 (Fig. 1), as a paraphyletic assemblage within Ornithopoda (e.g. Sereno 1999; Buchholz 2002; Butler 2005; Butler *et al.* 2008; Dieudonné *et al.* 2016, 2020; Han *et al.* 2018), with several

of the stratigraphically earliest ‘hypsilophodontids’ (e.g. *Agilisaurus*, *Hexinlusaurus*) placed outside Ornithopoda as non-cerapodan neornithischians.

Although a few older studies recovered a monophyletic ‘Hypsilophodontidae’ (e.g. Sereno 1986; Weishampel & Heinrich 1992), a study by Boyd *et al.* (2009), revising the taxonomy of *Thescelosaurus*, was the first to recover many of these taxa as a clade, but independent of *Hypsilophodon* and other more deeply nested ornithopods. This new clade included a dichotomy at its base, separating *Orodromeus*, *Oryctodromeus* and *Zephyrosaurus* (Orodrominae; Brown *et al.* 2013) from the Maastrichtian taxa *Parksosaurus* and *Thescelosaurus* (Thescelosaurinae; Brown & Druckenmiller 2011). The clade encompassing these two groups was later referred to as Thescelosauridae Sternberg, 1937 (see Brown *et al.* 2013), but its placement relative to Cerapoda remained somewhat ambiguous due to the absence of marginocephalian (Ceratopsia + Pachycephalosauria) operational taxonomic units (OTUs) from the analyses of Boyd *et al.* (2009) and Brown *et al.* (2013). Presenting the largest phylogenetic dataset of basal neornithischians compiled at the time (255 characters; 65 OTUs), Boyd (2015) recovered Thescelosauridae (incorrectly referred to as Parksosauridae in Boyd [2015]) as a sister group to Cerapoda (Fig. 1: H2 position; Fig. 2A). With respect to taxa usually regarded as either non-cerapodan neornithischians or ornithopod cerapodans, Boyd (2015) recovered: (1) *Stormbergia*, *Agilisaurus*, *Hexinlusaurus*, *Yandusaurus*, *Leaellynasaura*, *Yueosaurus*, *Jeholosaurus*, *Othnielosaurus* and the clade Thescelosauridae as non-cerapodan neornithischians; (2) *Koreanosaurus* as the sister-taxon to *Oryctodromeus*, within the thescelosaurid clade Orodrominae; (3) a clade of Gondwanan elasmarians (*Macrogryphosaurus*, *Notohypsilophodon*, *Talenkauen*) outside of Cerapoda within Thescelosauridae, as the sister clade to *Thescelosaurus*; (4) *Hypsilophodon* as the only non-iguanodontian ornithopod; and (5) a Gondwanan clade of early-diverging iguanodontians (*Atlascopcosaurus*, *Anabisetia*, *Gasparinisaura* and *Qantassaurus*). This matrix was later

updated by Madzia *et al.* (2018), adding a further 10 OTUs and modifying character scorings for 14 OTUs. Madzia *et al.* (2018) also recovered Thescelosauridae outside of Cerapoda, but with several key differences from the results of Boyd (2015) concerning the taxa of interest: (1) *Hypsilophodon* was placed outside of Cerapoda; (2) *Changchunsaurus* and *Haya* were recovered as orodromines rather than thescelosaurines, within Thescelosauridae; and (3) Elasmaria was placed inside Cerapoda as a clade of early-branching ornithopods, inclusive of the unnamed clade of Gondwanan ornithopods recovered in Boyd (2015).

[Insert FIGURE 2 near here]

Subsequent phylogenetic analyses have not recovered the 'hypsilophodontids' in a position congruent with Boyd (2015) and Madzia *et al.* (2018), except with the use of implied character weighting (Herne *et al.* 2019). To date, the largest morphological dataset of neornithischian taxa (380 characters; 72 OTUs) is the study presented by Han *et al.* (2018), based on their description of the early ceratopsian *Yinlong downsi* (Fig. 2B). This analysis recovered the 'hypsilophodontids' in a position congruent with H1, primarily within Ornithopoda. More specifically, among basal neornithischians and ornithopod cerapodans, Han *et al.* (2018) recovered: (1) *Agilisaurus* and *Hexinlusaurus* as the only non-cerapodan neornithischians; (2) *Haya*, *Jeholosaurus* and *Changchunsaurus* as a monophyletic group of Asian basal ornithopods (Jeholosauridae); and (3) *Koreanosaurus*, *Parksosaurus* and *Thescelosaurus* in a monophyletic group of non-iguanodontian ornithopods (?Thescelosaurinae). Most recently, Dieudonné *et al.* (2020) presented a new independent analysis of cerapodan relationships (342 characters; 72 OTUs) that also recovered the 'hypsilophodontids' within a position consistent with H1. Concerning basal neornithischians and ornithopod cerapodans, Dieudonné *et al.* (2020) recovered: (1) *Eocursor*, *Agilisaurus* and *Hexinlusaurus* as the only non-cerapodan neornithischians; (2) *Nanosaurus* as the earliest-branching ornithopod; (3) *Zephyrosaurus*, *Orodromeus* and *Koreanosaurus* as an early-

branching ornithopod clade, Orodrominae; (4) Parksosauridae within a Hypsilophodontidae clade together with *Hypsilophodon*; (5) Hypsilophodontidae, *Thescelosaurus*, and *Yueosaurus* + *Convolosaurus* as non-iguanodontian clypeodont ornithopods; and (6) Elasmaria within Iguanodontia.

Whereas matrices that produce a topology consistent with a H2 placement for the 'hypsilophodontids' (Boyd 2015; Madzia *et al.* 2018) seem to be the exception, recent descriptive papers of basal neornithischians, 'hypsilophodontids' and iguanodontians have used both H1 (e.g. Baron *et al.* 2017a; Salgado *et al.* 2017; Andrzejewski *et al.* 2018; Herne *et al.* 2018; Bell *et al.* 2019; Li *et al.* 2019; Yang *et al.* 2020) and H2 matrices (e.g. Herne *et al.* 2018; Madzia *et al.* 2018; Barta & Norell 2021), showing that both topologies are considered plausible within the literature and that there is no consensus on these relationships.

Methods

Data matrices

The matrices from Han *et al*. (2018) and Dieudonné *et al.* (2020) were chosen to represent hypothesis 1 (H1), while Boyd (2015) was used for hypothesis 2 (H2). While Dieudonné *et al.* (2020) is the most recent phylogenetic analysis to recover a topology congruent with H1, this analysis is contentious because it recovered heterodontosaurids within Pachycephalosauria, in contrast to many previous studies that identified the former as the earliest diverging ornithischian clade (e.g. Butler *et al*. 2008; Boyd 2015; Han *et al.* 2018; but see Xu *et al*. [2006] for an alternative view). However, testing hypotheses concerning the phylogenetic position of heterodontosaurids is beyond the scope of this study. To ensure that the results are not influenced by the differing positions of heterodontosaurids in Boyd (2015)

and Dieudonné *et al.* (2020), the data matrix from Han *et al.* (2018; which recovers heterodontosaurids as basal ornithischians, as in Boyd [2015]), was also analysed. The original data matrix presented in Boyd (2015) was selected to test H2 rather than the updated version of the matrix in Madzia *et al.* (2018) because the original Boyd (2015) matrix produced consensus trees with more resolved relationships for the taxa of interest. The matrix from Herne *et al.* (2019; a modified version of the Dieudonné *et al.* [2016] matrix) was not included in our analyses, despite recovering a H2 topology, because the H2 topology was only recovered using implied character weighting, which would make it an inappropriate comparator as all characters are weighted equally in the other analyses under consideration.

The data matrices from Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020) provide the basis for all of the analyses in this study. All three of these matrices were originally analysed using the traditional (heuristic) search option in TNT (Tree Analysis using New Technology: v. 1.5; Goloboff & Catalano 2016). The data matrices and original methods used by Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020) are summarized in Table 1. The original tree search from Boyd (2015) produced 36 most parsimonious trees (MPTs) and a relatively well-resolved strict consensus tree (Fig. 2A). The original tree searches from Han *et al.* (2018) and Dieudonné *et al.* (2020) produced 53,376 MPTs and 176 MPTs, respectively, both with relatively poorly resolved strict consensus trees, lacking resolution at the base of Cerapoda. Han *et al.* (2018) identified eight unstable taxa (*Albalophosaurus*, *Koreaceratops*, *Laquintasaura*, *Micropachycephalosaurus*, *Pisanosaurus*, *Yandusaurus*, *Yueosaurus* and *Zephyrosaurus*) that were subsequently removed using the 'Pruned trees' function in TNT to give a well-resolved reduced strict consensus tree (Fig. 2B). *Yandusaurus* was also identified as unstable in Dieudonné *et al.* (2020) and subsequently pruned to give a well-resolved reduced strict consensus (Fig. 2C). The original strict consensus trees from Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020) are

referred to hereafter as BSCT, HSCT and DSCT, respectively. Han *et al.* (2018) and Dieudonné *et al.* (2020) also produced and reported reduced consensus trees (described above), which are referred to in this study as HRSCT and DRSCT, respectively.

We performed an additional 92 phylogenetic analyses (Analyses A–K; Table 2) to test how character and taxon sampling may have influenced the relationships recovered by Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020). All taxa and character manipulation was implemented in TNT. Raw TNT files were downloaded from the supplementary material of Han *et al.* (2018) and Dieudonné *et al.* (2020). The supplement for Boyd (2015) only provides the character matrix as a .doc file, so scorings were inputted manually into Mesquite (v. 3.61, Maddison & Maddison 2019) and exported as a TNT file for our study. The TNT files from Boyd (2015), Han *et al*. (2018) and Dieudonné *et al.* (2020) are included in the supplement for this study (see Supplementary material 1–3). Analyses A–K were all conducted using a traditional search algorithm in TNT, with 10,000 replications and TBR swapping algorithm (holding 100 trees per replication). Strict consensus trees were generated and reported in all analyses. In some analyses, the TNT function prunnelsen or IterPCR (Pol & Escapa 2009) was used to identify unstable taxa that were pruned subsequently. When interpreting phylogenetic relationships derived from the re-analysis of matrices by Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020), we used the clade definitions as outlined in each of the original given studies (see Supplementary material 4, Table S1).

Stratigraphic congruence

All three matrices were re-analysed in TNT, replicating the original methods outlined in Table 1. The stratigraphic congruence of all of the MPTs, strict consensus trees and reduced strict consensus trees (if applicable) produced using the original methods of Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020) were assessed using the R-package `strap` (v.

1.4; Bell & Lloyd 2015), applying the command StratPhyloCongruence (with 1000 permutations). First and last appearance dates were retrieved from the Paleobiology Database (PBDB, https://paleobiodb.org/), checked against recent literature, and updated where necessary (see Supplementary material 4, Table S2). The as-yet undescribed North American orodromine 'Kaiparowits orodromine' from the Boyd (2015) matrix was given the full stratigraphic range of the Kaiparowits Formation (76.6–74.5 Ma; Roberts *et al*. 2013). Stratigraphic congruence was assessed using the following metrics: Stratigraphic Consistency Index (SCI: Huelsenbeck 1994); modified Manhattan Stratigraphic Measure (MSM*: Siddall 1998; Pol & Norell 2001), Gap Excess Ratio (GER: Wills 1999) and Minimum Implied Gap (MIG: Benton & Storrs 1994; Wills 1999). MIG was reported in millions of years. Time-calibrated strict and reduced strict consensus trees were produced using the R-packages `strap` (v. 1.4; Bell & Lloyd 2015) and `paleotree` (v. 3.3.25; Bapst & Wagner 2019).

Taxon sampling

Shared taxa. The matrix from Boyd (2015) shares 39 and 45 OTUs with the matrices from Han *et al*. (2018) and Dieudonné *et al*. (2020), respectively (Supplementary material 4, Table S1). To control for the effects of these studies using different OTUs the relationships among these shared OTUs were re-analysed using a traditional tree search in TNT, with 10,000 replications and TBR swapping algorithm (holding 100 trees per replication). OTUs that were not scored in matrices from Boyd (2015) and Han *et al*. (2018), or Boyd (2015) and Dieudonné *et al*. (2020) were removed prior to these tree searches. IterPCR was used to identify unstable taxa to prune them from the strict consensus trees.

Taxon completeness. In order to determine if taxon completeness was a factor in the differing relationships recovered, the OTUs coded by Boyd (2015), Han *et al.* (2018) and

Dieudonné *et al*. (2020) were assessed using an iteration of the Character Completeness Metric (CCM). The CCM was originally devised by Mannion & Upchurch (2010) as a method of assessing specimen completeness based on the proportion of phylogenetically informative characters preserved. Although initially developed to assess the sauropodomorph fossil record (Mannion & Upchurch 2010), it has since been used to quantify completeness in a range of extinct groups (e.g. Walther & Fröbisch 2013; Dean *et al.* 2016; Verrière *et al.* 2016; Tutin & Butler 2017).

Davies *et al.* (2017) used an alternative method to calculate CCM in fossil eutherian mammals, instead assessing the proportion of characters scored to a taxon relative to the total number of characters in the cladistic matrix. This is the iteration we use in this study to quantify completeness by taxon. Completeness was assessed in all three data matrices in R (v. 3.6.2) using the R-package `TreeTools` (v. 1; Smith 2019a) and `ape` (v. 5.4-1; Paradis *et al.* 2019). Unscored characters include unknown character states (?) and characters inapplicable to a taxon (-). The R-script used to compute completeness is provided in the supplement to this paper (see Supplementary material 5). The distribution of completeness scores across taxa from each character matrix was mapped onto the corresponding strict consensus tree using the R-packages `ape` (v. 5.4-1; Paradis *et al.* 2019), `phytools` (v. 0.7-47; Revell 2012) and `maps` (v. 3.3.0; Becker *et al.* 2016).

Tree searches of all three matrices were re-run excluding OTUs with varying low completeness: <10% complete (Analysis B.1); <20% (Analysis B.2); and <30% (Analysis B.3). All tree searches were undertaken in TNT using a traditional search with 10,000 replications and TBR swapping, holding 100 trees per replication (with a maximum of 100,000 trees held).

Taxon stability. Safe taxonomic reduction (Wilkinson 1995) identifies unstable OTUs that lack unique combinations of character scores, which can be removed *a priori* from a data matrix. This method reduces the number of MPTs generated and therefore increases the resolution of the resulting consensus trees, without affecting the inferred relationships of other OTUs in subsequent phylogenetic analyses. Safe taxonomic reduction was applied to all three data matrices using the R-package `Claddis` (v. 0.6.3; Lloyd 2016), applying the command safe_taxonomic_reduction. Unstable taxa identified using this method were removed from the respective data matrix subsequently using the taxcode- command in TNT. These data matrices were re-analysed using a traditional tree search in TNT, with 10,000 replications and TBR swapping algorithm (holding 100 trees per replication) as Analysis C.

IterPCR was used to identify unstable taxa to be pruned *a posteriori* from the strict consensus trees produced from tree searches of the Boyd (2015), Han *et al*. (2018) and Dieudonné *et al*. (2020) matrices (Analysis D). To identify their impact on resulting tree topologies, these unstable taxa were pruned from the strict consensus trees.

The positional congruence (reduced; PCR) of taxa from all three data matrices was identified using the *tcomp* function in TNT, following Pol & Escapa (2009). PCR values and completeness values of taxa from Boyd (2015), Han *et al*. (2018) and Dieudonné *et al.* (2020) were logged and tested for a significant correlation using Pearson's rank.

Character sampling

Shared characters. The character list from Boyd (2015) was manually compared with that from Han *et al.* (2018) and Dieudonné *et al*. (2020), respectively, to identify equivalent and shared characters in both analyses. These characters were also categorized based on whether they differed in character description and/or in the character states used. Characters that support the clades Thescelosauridae, Cerapoda or Ornithopoda in either the strict consensus

from Boyd (2015) or the reduced strict consensus from Han *et al*. (2018) and Dieudonné *et al*. (2020) were identified using the *apo* function in TNT. Characters found to support one of these clades in Boyd (2015), that are also characters identified as present in either Han *et al.* (2018) or Dieudonné *et al.* (2020), were mapped in TNT to identify whether their phylogenetic signals differ due to inconsistent scoring or taxon sampling.

Character ordering. The impact of ordering multi-state characters on tree resolution is debated within phylogenetics (Hauser & Presch 1991; Lipscomb 1992; Wilkinson 1992; Brazeau 2011; Grand *et al.* 2013; Rineau *et al.* 2015). All three of the data matrices analysed in this study include multi-state characters. Of these, 21 and four characters from Han *et al*. (2018) and Dieudonné *et al*. (2020), respectively, are ordered. No characters were ordered in Boyd (2015). Here we re-analysed the matrices from Han *et al.* (2018) and Dieudonné *et al.* (2020) in TNT with all characters unordered. This analysis (Analysis E) was implemented using a traditional tree search in TNT, with 10,000 replications and TBR swapping algorithm (holding 100 trees per replication).

Character partitions. In this study, characters from both matrices were grouped by: (1) skeletal element (e.g. femora); (2) into one of 12 skeletal regions (e.g. hind limb); and (3) as either craniodental or postcranial. The 12 skeletal regions (including seven craniodental and five postcranial sub-regions; Supplementary material 4, Fig. S1) are outlined in Supplementary material 4, Table S3. Characters not attributed to a specific skeletal element/region (e.g. H1, head shape in dorsal view), were excluded from subsequent analyses involving those partitions.

To identify the relative importance of characters from different skeletal regions, a series of 15 tree searches were run for each matrix, whereby characters from the following

partitions were excluded: (1) either all craniodental or all postcranial characters (Analysis F); (2) each of the 12 skeletal regions (Analysis G); and (3) characters pertaining to the most character-rich postcranial element from each data matrix (Analysis H). All tree searches were undertaken in TNT using a traditional search with 10,000 replications and TBR swapping, holding 100 trees per replication (with a maximum of 100,000 trees held). The *prunenelsen* and *pcrprune* TNT functions were used to identify unstable taxa to prune from the resulting strict consensus trees. Reduced strict consensus trees were reported when pruning resulted in increased resolution to taxa of interest, but without pruning over 10% of OTUs. Taxa were removed prior to the analysis if excluding characters from a specific partition in turn excluded all characters scored for that taxon.

Character completeness. The completeness of each character from all three matrices was assessed as the proportion of taxa scored per character. Completeness of characters was calculated using an R-script provided in the supplement to this paper. Character completeness was also compared between the 12 different skeletal regions used previously.

Tree searches of all three matrices were re-run excluding characters with varying low completeness to see if resulting topologies began to converge: ≤15% complete (Analysis I.1); ≤25 complete (Analysis I.2); and ≤35% complete (Analysis I.3). All tree searches were undertaken in TNT using a traditional search with 10,000 replications and TBR swapping, holding 100 trees per replication (with a maximum of 100,000 trees held).

Homoplasy. Data matrices from Boyd (2015), Han *et al*. (2018) and Dieudonné *et al.* (2020) were re-run using the original methods outlined above, but with implied weighting to see if their topologies converged. Implied weighting downweights characters based on how homoplastic the model considers them to be. The strength of this weighting is based on the

concavity constant, *k*, with lower values (e.g. *k*=1) more strongly downweighing homoplastic characters. The *piwe* function in TNT was used to activate implied weighting prior to the tree search with the following *k* values: 1 (Analysis J.1); 3 (Analysis J.2); and 5 (Analysis J.3). Character homoplasy was calculated from the strict consensus trees produced in the replicated tree searches of Boyd (2015), Han *et al.* (2018) and Dieudonné *et al*. (2020) using the *chomo* function in TNT. This function expresses homoplasy as the number of extra steps required to explain state changes of a character throughout a given tree topology as homologous rather than homoplastic. The relative distribution of homoplasy throughout different skeletal regions was presented as the percentage difference between the relative proportion of homoplasy (%) and relative proportion of total characters (%), for a given partition.

Tree searches were re-run for all three data matrices excluding characters of varying levels of homoplasy to see if the topologies would begin to converge. The cut-offs used were 95th percentile (Analysis K.1), 85th percentile (Analysis K.2) and 75th percentile (Analysis K.3). All tree searches were undertaken in TNT using a traditional search with 10,000 replications and TBR swapping, holding 100 trees per replication (with a maximum of 100,000 trees held).

Results

Stratigraphic congruence

The data matrices from Boyd (2015), Han *et al*. (2018) and Dieudonné *et al*. (2020) were re-analysed using their original search parameters (Table 1). The resulting consensus trees were also time-calibrated (Supplementary material 4, Fig. S2A–E). Tests of stratigraphic congruence show that the 36 MPTs from Boyd (2015) and 53,367 MPTs from Han *et al.* (2018) are generally equally congruent, while the 176 MPTs from Dieudonné *et al*. (2020)

(Table 3) are the least congruent. The strict consensus tree from Boyd (2015) (Supplementary material 4, Fig. S2A) is more congruent than that from Han *et al.* (2018) (Supplementary material 4, Fig. S2B); however, the reduced strict consensus from Han *et al.* (2018) (Supplementary material 4, Fig. S2C) is generally more congruent than the strict consensus from Boyd (2015) (Table 4). The strict consensus tree from Boyd (2015) is more stratigraphically congruent than both the strict and reduced strict consensus trees from Dieudonné *et al.* (2020; Table 4; compare Supplementary material 4, Fig. S2D, E).

Taxon sampling

Shared taxa. Including only the 39 taxa shared by both Boyd (2015; 60% overlap) and Han *et al*. (2018; 54.2% overlap) in a tree search of the pruned Boyd (2015) matrix results in 66 MPTs and a poorly resolved strict consensus tree (Supplementary material 4, Fig. S3A; Analysis A.1). However, two unstable taxa (*Yueosaurus* and *Wannanosaurus*) were identified and pruned subsequently, giving a well-resolved reduced strict consensus tree (Fig. 3A) that possesses a number of differences from BSCT: (1) heterodontosaurids were resolved as basal neornithischians; (2) Thescelosauridae was found within Ornithopoda; (3) *Changchunsaurus* + Thescelosauridae were found as the sister group of Clypeodonta (*Hypsilophodon* + Iguanodontia); and (4) *Jeholosaurus*, *Agilisaurus*, *Yandusaurus* and *Hexinlusaurus* formed a clade within Thescelosauridae.

Removing taxa from the matrix by Han *et al.* (2018) that are not scored in Boyd (2015) produced 164 MPTs and a poorly resolved strict consensus tree (Supplementary material 4, Fig. S3B). Following pruning of three unstable taxa (*Abrictosaurus*, *Yueosaurus* and *Micropachycephalosaurus*) a well-resolved reduced strict consensus tree was obtained (Fig. 3B), which recovered several differences in topology from HRSCT: (1) heterodontosaurids are found within Cerapoda in a sister-group relationship with

Marginocephalia; and (2) *Eocursor* and *Lesothosaurus* are recovered within Neornithischia, but in earlier-branching positions than *Agilisaurus* and *Hexinlusaurus*; (3) *Yandusaurus* is identified as the sister group of Cerapoda; and (4) *Orodromeus* forms a clade with *Zephyrosaurus* (Orodrominae *sensu* Brown *et al.* 2013).

[Insert FIGURE 3 near here]

Including only the 45 taxa coded in both Boyd (2015; 69.2% overlap) and Dieudonné *et al*. (2020; 62.5% overlap) when conducting a tree search of the Boyd (2015) matrix, produces 180 MPTs and a poorly resolved strict consensus (Supplementary material 4, Fig. S4A; Analysis A.2). IterPCR identified six unstable taxa (*Changchunsaurus*, *Hypsilophodon*, *Macrogryphosaurus*, *Talenkauen*, *Valdosaurus*, *Wannanosaurus*) and pruning these resulted in a well-resolved reduced strict consensus (Fig. 4A) that shows several differences compared with BSCT: (1) heterodontosaurids recovered as basal neornithischians; (2) *Stenopelix* found to be a non-cerapodan neornithischian; (3) Thescelosauridae nested within Ornithopoda in a sister-group relationship with Iguanodontia; (4) *Jeholosaurus*, *Agilisaurus*, *Yandusaurus* and *Hexinlusaurus* found within Thescelosauridae; and (5) *Gasparinisaura* recovered as an iguanodontian. Excluding taxa from Dieudonné *et al*. (2020) that are not coded in Boyd (2015) results in two MPTs and a well resolved strict consensus (Fig. 4B; Supplementary material 4, Fig. S4B). This consensus differs from DSCT in recovering: (1) *Eocursor* as a basal ornithischian; (2) *Yandusaurus* within Ornithopoda; (3) Parksosauridae + Hypsilophodontidae in a polytomy at the base of Clypeodonta; and (4) Elasmaria outside of Iguanodontia and forming a clade with *Yueosaurus*.

[Insert FIGURE 4 near here]

Taxon sampling is shown here to substantially alter the implied relationships recovered across all three matrices. Boyd (2015) and Han *et al.* (2018) are particularly impacted by taxon inclusion, with numerous changes identified concerning the taxa of

interest and clades that are frequently considered to have ambiguous affinities (e.g. heterodontosaurids, thescelosaurids).

Taxon completeness. Taxa coded in the Boyd (2015) matrix are generally less complete (average = 41.1%) than those scored in both Han *et al*. (2018; average = 50.9%) and Dieudonné *et al.* (2020; average = 46.8%) (Fig. 5).

[Insert FIGURE 5 near here]

Neornithischian taxa with very low completeness in Boyd (2015) are found throughout the tree (Supplementary material 4, Fig. S5A) and tend to fall within polytomies (e.g. *Yandusaurus*, *Macrogryphosaurus*, *Notohypsilophodon*, *Micropachycephalosaurus*, *Stenopelix*, *Callovosaurus*). The few highly complete taxa in Boyd (2015) are consistently recovered in resolved positions (e.g. *Dryosaurus*, *Dysalotosaurus*, *Hypsilophodon*, *Orodromeus*).

In the strict consensus tree from Han *et al.* (2018; Supplementary material 4, Fig. S5B), the polytomy at the base of Cerapoda contains both highly incomplete (e.g. *Micropachycephalosaurus*, *Yueosaurus*, *Koreanosaurus*) and highly complete taxa (e.g. *Orodromeus*, *Jeholosaurus*, *Hypsilophodon*, *Haya*). All eight of the taxa pruned in HRSCT have <25% completeness.

The polytomy at the base of Cerapoda in the strict consensus tree from Dieudonné *et al*. (2020; Supplementary material 4, Fig. S5C) contains moderately complete taxa (>50%), with the exception of *Yandusaurus* (22%), which was pruned from the original analysis after being identified as unstable. Taxa of moderate completeness generally occur throughout the tree, although areas within Iguanodontia in particular have taxa that are both highly incomplete (e.g. the elasmarians, which are recovered as a polytomy) and highly complete (e.g. *Tenontosaurus*, *Dryosaurus*, *Camptosaurus*, *Iguanodon*).

Removing the most incomplete OTUs from Boyd (2015) reduces the resolution of inferred relationships for the taxa of interest in this study compared to BSCT (Analyses B1–3.B2015; Supplementary material 4, Figs S6–8). The thescelosaurines are particularly affected, falling as a polytomy at the base of Thescelosauridae when OTUs <30% complete (n=28) are removed prior to the tree search (Analysis B.3.B2015; Supplementary material 4, Fig. S8). Removing highly incomplete taxa from Boyd (2015) also recovers *Jeholosaurus* within Thescelosauridae, rather than as a non-thescelosaurid neornithischian as in BSCT. Excluding OTUs <10% complete prior to the tree search of Han *et al.* (2018; Analysis B.1.H2018; Supplementary material 4, Fig. S9) and Dieudonné *et al.* (2020; Analysis B.1.D2020; Supplementary material 4, Fig. S12) does not impact the resulting tree topology, perhaps due to the small number of highly incomplete taxa in both matrices (n=3 and n=2, respectively; compared to n=7 in Boyd [2015]). Removing taxa <20% complete from Han *et al.* (2018; Analysis B.2.H2018; Supplementary material 4, Fig. S10) largely improves the resolution of implied relationships, but further removal of incomplete taxa (<30% complete; Analysis B.3.H2018; Supplementary material 4, Fig. S11) fully resolves the polytomy at the base of Cerapoda. Analysis B.2-3.H2018 (Supplementary material 4, Figs S10, S11) also recovers *Lesothosaurus* within Thyreophora (consistent with Butler *et al.* [2008] and Boyd [2015]), rather than as a basal genasaurian, as in HSCT. Removing taxa <20% complete from Dieudonné *et al.* (2020; Analysis B.2.D2020; Supplementary material 4, Fig. S13) resolves the polytomy at the base of Cerapoda in DSCT, but many 'hypsilophodontids' remain poorly resolved within Ornithopoda. Further excluding taxa <30% complete (Analysis B.3.D2020, Figure S14) resolves the relationships of the remaining 'hypsilophodontids', but other groups (e.g. thyreophorans) become less resolved.

Taxa from Boyd (2015) are generally more incomplete than those from Han *et al.* (2018) and Dieudonné *et al.* (2020). Removing incomplete taxa from Boyd (2015) reduces

tree resolution, while the same action increases tree resolution for the Dieudonné *et al.* (2020) and, especially, the Han *et al*. (2018) datasets. Relationships of the taxa of interest generally remain consistent with those presented in the original studies.

Taxon Stability. Application of safe taxonomic reduction to the Boyd (2015) matrix identified *Thescelosaurus garbanii* as taxonomically equivalent to *Thescelosaurus neglectus*. *Th. garbanii* was removed from the Boyd (2015) matrix prior to re-analysis, which generated 36 MPTs with a well-resolved strict consensus tree (Supplementary material 4, Fig. S15; Analysis C.B2015). Apart from the absence of *Th. garbanii*, this strict consensus tree is identical to BSCT. When applied to the matrix of Han *et al.* (2018), *Yueosaurus tiantaiensis* was identified as taxonomically equivalent to *Hypsilophodon foxii* and *Orodromeus makelai*. *Y. tiantaiensis* was removed from the matrix, which was then re-analysed in TNT, producing 3,439 MPTs and a poorly-resolved strict consensus tree (Supplementary material 4, Fig. S16; Analysis C.H2018) identical to HSCT, except in recovering *Jeholosaurus* as the sister-taxon to *Changchunsaurus*. No OTUs were identified as safe to remove from the Dieudonné *et al.* (2020) dataset.

IterPCR identified *Lycorhinus* as an unstable taxon that could be pruned to improve the resolution of the strict consensus tree generated from the Boyd (2015) matrix. When this taxon is pruned from the strict consensus tree (Analysis D.1.B2015; Supplementary material 4, Fig. S17) the relationships of heterodontosaurids and thescelosaurines are better resolved. Excluding *Lycorhinus* prior to the tree search produces 12 MPTs with a strict consensus (Analysis D.2.B2015; Supplementary material 4, Fig. S18) that is equally well-resolved as that resulting from Analysis D.2.B2015, but with several minor branch swaps.

Eight unstable taxa (*Yueosaurus*, *Pisanosaurus*, *Yandusaurus*, *Zephyrosaurus*, *Albalophosaurus*, *Laquintasaura*, *Micropachycephalosaurus* and *Koreaceratops*) were

pruned in the original reduced strict consensus of 53,376 MPTs derived from Han *et al*. (2018). Here, IterPCR identifies nine taxa (*Yueosaurus*, *Pisanosaurus*, *Yandusaurus*, *Zephyrosaurus*, *Albalophosaurus*, *Laquintasaura*, *Micropachycephalosaurus*, *Aquilops* and *Yamaceratops*) that could be pruned to improve the tree resolution; however, the additional pruning of *Aquilops* and *Yamaceratops* only increases the resolution among neoceratopsian taxa (Analysis D.1.H2018; Supplementary material 4, Fig. S19). These nine taxa were also excluded prior to the tree search replication, producing 660 MPTs (Analysis D.2.H2018). The strict consensus of these trees is relatively well-resolved (Supplementary material 4, Fig. S20), similar to HRSCT, but comparably lacking in resolution among the heterodontosaurids, basal neornithischians and basal ceratopsians.

Dieudonné *et al.* (2020) pruned *Yandusaurus* to increase the resolution of the original strict consensus of 176 MPTs. Here, IterPCR identified *Macrogryphosaurus*, *Mahuidacursor* and *Yandusaurus* as unstable in the tree search. Pruning these three taxa (Analysis D.1.D2020; Supplementary material 4, Fig. S21) resolves the polytomies at the base of Cerapoda and Elasmaria recovered in the original strict consensus from Dieudonné *et al.* (2020). Excluding these three taxa prior to the tree search replication produces 40 MPTs and a strict consensus tree (Supplementary material 4, Fig. S22) similar to Figure S21, but lacking resolution among the iguanodontians (Analysis D.2.D2020).

Taxa from the 36 MPTs generated from the Boyd (2015) matrix are generally far more positionally congruent than those from the 53,376 MPTs and 176 MPTs generated from the matrices by Han *et al*. (2018) and Dieudonné *et al.* (2020), respectively (Supplementary material 4, Fig. S23). The least positionally congruent taxa from Boyd (2015) are *Hexinlusaurus*, *Leaellynasaura* and *Yandusaurus* (PCR = 0.958), which are all recovered in a polytomy at the base of Neornithischia in the strict consensus tree (Fig. 2A). In the data matrix from Han *et al*. (2018), there are seven taxa that are outliers in their poor positional

congruence, with PCR values ranging from 0.557–0.784 (*Pisanosaurus*, *Laquintasaura*, *Micropachycephalosaurus*, *Yueosaurus*, *Yandusaurus*, *Lesothosaurus* and *Eocursor*). All of these taxa (except for *Lesothosaurus*) have low completeness scores (<30%). Taxa from Dieudonné *et al.* (2020) are highly positionally congruent (PCR = >0.976), with the exception of one outlier taxon, *Yandusaurus*, which has a PCR value of 0.596. PCR values and completeness are not correlated significantly in taxa from either Boyd (2015) or Dieudonné *et al.* (2020; Table 5). PCR values and completeness are significantly positively correlated in taxa from Han *et al.* (2018), but do not significantly correlate when those seven PCR-outlier taxa are excluded.

Safe taxonomic reduction has little impact on tree resolution in Analysis C, but at least in Han *et al.* (2018) *a priori* exclusion of taxa identified as safe to remove significantly reduced the number of MPTs generated from 53,376 to 3,439, which should improve consensus tree resolution. Following a tree search, IterPCR finds unstable taxa in each of the three matrices, although considerably more in Han *et al.* (2018). When these unstable taxa are pruned following a tree search analysis, the resulting reduced strict consensus trees generated are generally better resolved than if those taxa were excluded *a priori*. IterPCR identifies *Yandusaurus* as unstable in both Han *et al.* (2018) and Dieudonné *et al.* (2020). *Yandusaurus* is consistently identified as unstable across all three matrices, being recovered in a polytomy in all three original strict consensus trees and identified as having low PCR values. Taxa from Han *et al.* (2018) are less positionally congruent than those from Boyd (2015) and Dieudonné *et al.* (2020). Generally, there is no relationship between completeness and positional congruence (although the two values significantly correlate in Han *et al.* [2018]), which is particularly evident in several taxa that are highly incomplete in two or all of the matrices investigated but have vastly different PCR values following tree searches (e.g. *Micropachycephalosaurus*, *Yueosaurus*).

Character sampling

Shared characters. The matrix from Boyd (2015) shares 79 characters with Han *et al.* (2018); however, only 38 of those shared characters use the same character states (including those where the states are reversed) and thus capture the same morphological variation (see Table S4). Boyd (2015) shares more characters with Dieudonné *et al.* (2020; n=118), 58 of which use the same character states (Supplementary material 4, Table S4).

The clade Thescelosauridae in Boyd (2015) is supported by state changes in seven characters. None of these characters are coded in Han *et al.* (2018), but five are coded in Dieudonné *et al.* (2020). Of these, only character B255 (=D21; unfused/fused premaxillae) uses the same character states in both matrices and are thus equivalent. In addition, there is no conflict in the coding of this character between the two matrices. In several of the remaining shared characters, the state partitioning has been subtly reworked in Dieudonné *et al.* (2020) (e.g. B98=D122; angle between the base and long axis of the braincase), which has in turn caused substantial changes to how these characters have been coded and therefore how they infer evolutionary change.

Cerapoda is supported by state changes in seven characters in Boyd (2015). Three of these characters are also coded in Han *et al.* (2018), and three are coded in Dieudonné *et al.* (2020). B123 (distribution of enamel on maxillary and dentary teeth) is found in both Han *et al.* (2018; H226) and Dieudonné *et al.* (2020; D171). D171 is equivalent to B123 (although uses slightly different wording) but is not supportive of Cerapoda in Dieudonné *et al.* (2020). H226 is not equivalent to B123 as it has an extra state, but both H226:1 and B123:1 (asymmetrical distribution of enamel on maxillary and dentary teeth) are supportive of Cerapoda in their respective phylogenetic analyses.

In Boyd (2015), the clade Ornithopoda is supported by state changes in seven characters. Two of these characters are also coded in the matrix by Han *et al.* (2018), while three are coded in Dieudonné *et al.* (2020). Two of these shared characters are equivalent to each other (B28=H44, maxilla/jugal suture; B250=D219, absence/presence of ossified hypaxial tendons along the tail); however, neither are supportive of Ornithopoda in their respective phylogenetic analyses.

Character ordering. Unordering all characters in Han *et al.* (2018) (Analysis E.H2018) generated 47,801 MPTs. The strict consensus of these MPTs (Supplementary material 4, Fig. S24) is slightly more poorly resolved than when 21 characters are ordered, particularly in the resolution of the ceratopsians.

Unordering all characters in Dieudonné *et al.* (2020) (Analysis E.D2020) produced 100 MPTs and a relatively poorly resolved strict consensus tree (Supplementary material 4, Fig. S25). This strict consensus recovers several 'hypsilophodontids' (*Koreanosaurus, Kulindadromeus, Orodromeus, Parksosaurus, Yueosaurus, Zephyrosaurus*), originally recovered in resolved positions within Ornithopoda, in a clade at the collapsed base of Cerapoda. These results show how even minimal use of character ordering (e.g. the four ordered characters in Dieudonné *et al.* [2020]) can have an impact on the resulting tree topology and inferred relationships of taxa.

Character partitions. As is typical for many morphological matrices (Mounce *et al.* 2016), craniodental characters are more abundantly represented than postcranial characters in Boyd (2015; 55.7%), Han *et al.* (2018; 59.5%) and Dieudonné *et al.* (2020; 56.4%) (Table 6). Of these craniodental characters, those relating to the dentition make up a larger proportion of the total characters in Boyd (2015; 12.2%) than in Han *et al.* (2018; 8.4%) and Dieudonné *et*

al. (2020; 10.5%). In Boyd (2015) and Dieudonné *et al.* (2020), the hind limb region contains the highest proportion of characters from the postcranium (16.5% and 12.9%, respectively; compared to 8.7% in Han *et al.* 2018). The pelvic girdle is the best represented postcranial region in the matrix of Han *et al.* (2018), comprising 11.3% of all characters, compared to 10.2% and 10.8% in Boyd (2015) and Dieudonné *et al.* (2020), respectively. Of the main postcranial regions, the pectoral girdle is the most poorly represented in all three matrices (Boyd 2015, 3.1%; Han *et al.* 2018, 3.2%; Dieudonné *et al.* 2020, 3.5%). The most character-rich postcranial element in Boyd (2015) is the femur, making up 7.8% of all characters (for comparison: Han *et al.* 2018, 3.2%; Dieudonné *et al.* 2020, 5%). In both Han *et al.* (2018) and Dieudonné *et al.* (2020) the ilium is the most character-rich postcranial element, making up 5.5% and 6.1% of total characters, respectively (compared to 4.7% in Boyd 2015).

Excluding all craniodental characters (n=142) from Boyd (2015; Analysis F.1.B2015) produced 33,800 MPTs and a poorly resolved strict consensus (Supplementary material 4, Fig. S26) with a polytomy of cerapodan and some basal neornithischian and ornithischian taxa. Removing postcranial characters (n=113; Analysis F.2.B2015) produced 6,800 MPTs with a strict consensus tree consisting of a large polytomy of all ingroup taxa (Supplementary material 4, Fig. S27). Excluding craniodental characters (n=226) from the analysis by Han *et al*. (2018; Analysis F.1.H2018) produced 44,600 MPTs and a polytomous strict consensus tree (Supplementary material 4, Fig. S28), while removing postcranial characters (n=154; Analysis F.2.H2018) produced 600 MPTs with a strict consensus consisting of a large polytomy at the base of Neornithischia (Supplementary material 4, Fig. S29). Excluding craniodental characters (n=193) from Dieudonné *et al.* (2020; Analysis F.1.D2020) results in 6,800 MPTs and a strict consensus consisting of a large polytomy of all ingroup taxa (Figure S30). When postcranial characters were excluded (n=149) prior to the reanalysis of Dieudonné *et al.* (2020; Analysis F.2.H2018), 16,900 MPTs were produced, with a strict

consensus tree characterized primarily by a very large polytomy (Supplementary material 4, Fig. S31).

When excluding character partitions pertaining to different skeletal regions from Boyd (2015; Analysis G.1-12.B2015; Supplementary material 4, Figs S32–42), Thescelosauridae remained in a non-cerapodan neornithischian position in only five of the 11 analyses (character partitions excluded: braincase; mandible; vertebral column; pectoral girdle; pelvic girdle). When either all dental (n=31) characters or all hind limb (n=42) characters were excluded from the Boyd (2015) matrix, Thescelosauridae was recovered nested within Ornithopoda (Supplementary material 4, Figs S37B, S42A, B). In Han *et al*. (2018), the 'hypsilophodontids' remain within Cerapoda when each of the 12 partitions are excluded (Analysis G.1-12.H2018; Supplementary material 4, Figs S43–54). Most of the strict consensus trees produced following exclusion of character partitions from Han *et al.* (2018) present topologies and inferred relationships very similar to those in the original strict consensus (rostral; palatal; cranial roof; skull base; mandibular; dental; vertebral column; pectoral girdle; forelimb). However, excluding some of the character partitions from Han *et al*. (2018) recovers controversial placements for clades, such as the inclusion of heterodontosaurids within Cerapoda (braincase; hind limb). 'Hypsilophodontids' also remained within Cerapoda when each of the 12 character partitions were excluded from the Dieudonné *et al.* (2020) matrix (Analysis G.1-12.D2020; Supplementary material 4, Figs S55–66).

Removing femoral characters (the most character-rich postcranial element, n=20) from the analysis by Boyd (2015), resulted in 1,700 MPTs and the recovery of Thescelosauridae within Ornithopoda following pruning of six unstable taxa from the strict consensus (Analysis H.B2015; Supplementary material 4, Fig. S67B). Femoral characters B212 (presence/absence of a trench between the greater trochanter and the head of the

femur), B213 (convex/flat lateral surface of the greater trochanter) and B223 (form of the anterior intercondylar groove) were identified as particularly problematic in their effects on tree topology. Excluding these three characters from the Boyd (2015) data matrix prior to tree searching produces 600 MPTs and a relatively well-resolved strict consensus tree (Fig. 6), recovering Thescelosauridae within Ornithopoda. Excluding iliac characters from the Han *et al*. (2018, n=21) matrix resulted in 2,696 MPTS and a poorly-resolved strict consensus tree (Analysis H.H2018; Supplementary material 4, Fig. S68). This strict consensus is largely the same as HSCT, but removing iliac characters resolves the polytomy at the base of Ornithischia. Removing iliac characters from Dieudonné *et al.* (2020, n=21) prior to the tree search recovers the clades Orodrominae, Marginocephalia and Clypeodonta in a polytomy at the collapsed base of Neornithischia (Analysis H.D2020; Supplementary material 4, Fig. S69).

[Insert FIGURE 6 near here]

Excluding either cranial or postcranial character partitions from each of the three analyses produced consensus trees with poor resolution. Removing craniodental characters from Boyd (2015) had less negative impact on the resulting tree resolution than removing postcranial characters, in contrast to the results from Han *et al.* (2018) and Dieudonné *et al.* (2018) where the opposite was found. Excluding specific character regions affected the outputs of the three matrices differently, but seemed to impact the inferred relationships of Dieudonné *et al.* (2020) the least. Removing certain skeletal partitions from Boyd (2015) and Han *et al.* (2018) recovered controversial placements for some taxa (e.g. thescelosaurids in Boyd [2015]; heterodontosaurids in Han *et al.* [2018]), differing significantly from the results of their original analyses.

Character completeness. Character completeness scores are generally highest in Han *et al.* (2018; average = 50.9%), followed by Dieudonné *et al.* (2020; average = 46.7%) and lastly, Boyd (2015; average = 41.1%) (Fig. 7). In Boyd (2015), postcranial characters (average = 43.9%) have higher completeness scores than cranial characters (average = 38.7%). In contrast, cranial character completeness scores (average = 54.6%) were far higher than those for postcranial characters (average = 45%) in Han *et al.* (2018), while both partitions had equal completeness in Dieudonné *et al.* (2020; craniodental average = 46.2%; postcranial average = 46.7%).

[Insert FIGURE 7 near here]

Dental characters have the highest completeness scores of any craniodental partition for each of the three matrices (Supplementary material 4, Fig. S70A). Palatal characters comprise the cranial partition with the lowest completeness scores in Han *et al.* (2018) and Dieudonné *et al.* (2020), whereas in Boyd (2015) braincase character scores are lowest (although no palatal characters are scored in Boyd [2015]). In the matrix of Boyd (2015), hind limb elements exhibit the highest character completeness scores of all postcranial regions (Supplementary material 4, Fig. S70B). Pelvic characters are generally the most complete postcranial partition in Han *et al.* (2018) and Dieudonné *et al.* (2020). The forelimb region has the lowest character completeness scores across all three matrices.

Excluding characters with scores of less than 25% (Analysis I.1-2.B2015; Supplementary material 4, Figs S71, S72) from the Boyd (2015) matrix prior to the tree search produces a well resolved strict consensus with tree topologies very similar to the original strict consensus from Boyd (2015). Excluding characters whose scoring is ≤35% complete (n=93) prior to the tree search of Boyd (2015; Analysis I.3.B2015) results in 1,300 MPTs and a poorly resolved strict consensus tree with the contents of Thescelosauridae and Cerapoda in a large polytomy (Supplementary material 4, Fig. S73).

When characters that were scored for less than 15% (n=6) and 25% (n=22) were excluded from Han *et al.* (2018) prior to the tree search, 37,644 and 14,015 MPTs were produced, respectively, both with poorly resolved strict consensus trees (Supplementary material 4, Figs S74, S75) very similar to the original analysis from Han *et al.* (2018). Further excluding characters scored with up to 35% completeness (n=68) produced 5,361 MPTs and a similar poorly resolved strict consensus tree (Supplementary material 4, Fig. S76), but with slightly better resolution among basal ornithischian taxa.

Excluding characters whose scoring is less than 15% complete (n=7) from Dieudonné *et al.* (2020; Analysis I.1.D2020) produced 1,780 MPTs and a poorly resolved strict consensus tree (Supplementary material 4, Fig. S77) very similar to DSCT, but with a polytomy of basal ornithopods. When further characters with scores that are under 25% (n=22) and 35% (n=74) complete are excluded from Dieudonné *et al*. (2020), 4,559 and 1,981 MPTs are produced, respectively, both with much more poorly resolved strict consensus trees (Supplementary material 4, Figs S78, S79).

These results demonstrate that removing characters with low completeness scores has a negative effect on tree resolution in all three matrices, but does not result in major changes to inferred relationships.

Homoplasy. The tree search of the Boyd (2015) matrix with an implied weighting of k=5 (Analysis J.1.B2015), produced 538 MPTs with a relatively well-resolved strict consensus tree (Supplementary material 4, Fig. S80) that showed several topological differences to the original Boyd (2015) strict consensus tree produced where all characters were equally weighted: (1) the ‘elasmarian thescelosaurids’ *Notohypsilophodon*, *Talenkauen* and *Macrogryphosaurus* are recovered within Cerapoda as ornithopods; (2) the basal marginocephalian *Stenopelix* is placed as a basal neornithischian; (3) *Hypsilophodon* and

Micropachycephalosaurus are recovered outside of Cerapoda as basal thescelosaurids; and (4) the Asian taxa *Yueosaurus* and *Jeholosaurus* are recovered within Thescelosauridae. As the concavity constant is reduced (*k*=3 and *k*=1; Analysis J.2–3.B2015) the resolution of the resulting tree topologies decreases, with larger polytomies at the base of Cerapoda, but the contents of Thescelosauridae and Cerapoda remain the same.

A tree search of the matrix by Han *et al.* (2018) with implied weighting using a high concavity constant (*k*=5) produced 384 MPTs with a poorly resolved strict consensus tree (Supplementary material 4, Fig. S21). This strict consensus showed very few differences to that produced when all characters were weighted equally, but is generally better resolved: (1) the polytomy at the base of Ornithischia is resolved; (2) Marginocephalia (Ceratopsia + Pachycephalosauria) is recovered; (3) iguanodontian relationships are fully resolved; (4) *Yandusaurus* is recovered as an early-branching pachycephalosaurid; and (5) the possible ceratopsian *Albalophosaurus* is placed in a polytomy at the base of Genasauria. Reducing the concavity constant to *k*=3 does not impact the topology of the strict consensus tree, and further reducing it to *k*=1 only results in the movement of *Stenopelix* to the base of Marginocephalia.

When an implied weighting of *k*=5 was applied prior to a tree search of the Dieudonné *et al.* (2020; Analysis J.1.D2020) matrix, three MPTs were produced with a well-resolved strict consensus tree (Supplementary material 4, Fig. S86). This strict consensus tree showed several differences concerning 'hypsilophodontid' taxa to that of the DSCT, recovering: (1) *Hexinlusaurus* within Ornithopoda in a sister-taxa relationship with *Kulindadromeus*; (2) *Haya* and *Changchunsaurus* + *Jeholosaurus* in an Asian clade 'Jeholosauridae' *sensu* Han *et al.* (2012); (3) 'Jeholosauridae' in a sister-group relationship with Orodrominae; (4) *Yandusaurus* in a resolved position within Ornithopoda, as a sister-group to *Gasparinisaura* + *Parksosaurus*; and (5) *Gasparinisaura* + *Parksosaurus* basal to

Hypsilophodon and outside of Clypeodonta and Hypsilophodontidae. Reducing the concavity constant to *k*=3 (Analysis J.2.D2020) does not significantly impact the relationships of the strict consensus tree (Supplementary material 4, Fig. S87), but further reduction to *k*=1 (Analysis J.3.D2020) recovers controversial relationships: Marginocephalia is not recovered and Pachycephalosauria falls out as an early branching ornithischian clade (Supplementary material 4, Fig. S88).

Characters from Dieudonné *et al.* (2020) are far more homoplastic than those from both Boyd (2015) and Han *et al.* (2018) (Supplementary material 4, Fig. S89). The distribution of homoplasy throughout the skeleton differs among all three matrices (Supplementary material 4, Fig. S90). In Boyd (2015), pelvic girdle characters are the most homoplastic relative to their weighting. Dental characters are highly homoplastic across all three matrices, but especially so in Han *et al.* (2018) relative to the rest of the skeleton. The most homoplastic dental characters in Boyd (2015; B137) and Dieudonné *et al.* (2020; D150) concern the number of dentary teeth, while in Han *et al.* (2018) the most homoplastic character relates to the number of maxillary teeth (H203). For Han *et al.* (2018) and Dieudonné *et al.* (2020) these dental characters are also the most homoplastic of all characters in the given matrix. In all three matrices, rostral characters are the least homoplastic.

Excluding characters with varying degrees of homoplasy did not cause the topologies to converge. Excluding homoplastic characters in each of the three ranges (>95th, >85th, >75th percentile) investigated from Boyd (2015; Analysis K.1-3.B2015) generates highly unresolved strict consensus trees consisting of a polytomy at the base of Ornithischia (Supplementary material 4, Figs S91–93). Excluding characters in the top 95th percentile from the matrix by Han *et al.* (2018; Analysis K.1.H2018) generated 44,900 MPTs and a poorly resolved strict consensus with a polytomy at the base of Ornithischia, but recovered

the clades Thyreophora, Iguanodontia and Neoceratopsia (Supplementary material 4, Fig. S94). Exclusion of homoplastic characters in the top 85th percentile (Analysis K.2.H2018) produced 40,500 MPTs with a poorly resolved strict consensus (Supplementary material 4, Fig. S95) very similar to HSCT, but is better resolved, recovering Marginocephalia and Genasauria. Further exclusion of homoplastic characters (Analysis K.3.H2018) generates a highly unresolved strict consensus from 99,100 MPTs consisting of a polytomy at the base of Ornithischia (Supplementary material 4, Fig. S96). Excluding characters with homoplasy over the 95th percentile from Dieudonné *et al.* (2020; Analysis K.1.D2020) resulted in 3,779 MPTs and a poorly resolved strict consensus tree (Supplementary material 4, Fig. S97), similar to DSCT, but with a polytomy at the base of Ornithopoda and poorer resolution among pachycephalosaurians. Further exclusion of homoplastic characters (Analysis K.2–3.D2020) produces an even more poorly resolved strict consensus (Supplementary material 4, Fig. S98–99), with a polytomy at the base of Ornithischia and Neornithischia.

Use of implied weighting does not impact the topology of Han *et al.* (2018), while in Dieudonné *et al.* (2020) only the strongest use of implied weighting ($k = 1$) causes significant topological changes to the ornithischian tree (e.g. recovering pachycephalosaurians at the base of Ornithischia). The use of implied weighting in the Boyd (2015) matrix results in a strict consensus that is less well resolved than BSCT and contains several major changes in topology concerning the taxa of interest. Excluding highly homoplastic characters from all three matrices does not alter the topologies or relationships observed from the original studies but does significantly reduce the resolution of the consensus trees produced.

Discussion

Taxon sampling

Studies investigating phylogenetic methods agree overwhelmingly that more extensive taxon sampling, including the addition of highly incomplete taxa, is beneficial to analyses (e.g. Kearney 2002; Zwickl & Hillis 2002; Kearney & Clark 2003; Wiens 2006; Butler & Upchurch 2007; Heath *et al.* 2008; Prevosti & Chemisquy 2010; Wiens & Morrill 2011; but see Rosenberg & Kumar [2001] for an alternative view). However, it is also clear that unstable taxa can have a negative effect on tree resolution and obscure inferred relationships, with reduced consensus methods being utilised in several recent analyses of ornithischian relationships (e.g. Maidment 2010; Thompson *et al.* 2012; Baron *et al.* 2017a; Han *et al.* 2018; Dieudonné *et al.* 2020).

In addition, we find here that the choice of taxa included within an analysis can significantly impact the inferred relationships, particularly concerning 'wildcard' groups with controversial or disputed phylogenetic placements (e.g. heterodontosaurids, thescelosaurids) (Brusatte 2010; Raven & Maidment 2018). Although there is a reasonable amount of overlap between taxa in Boyd (2015) and either of the two H1 matrices (>60% overlap), removing unique taxa in Analysis A resulted in several substantial changes from the original reported tree topologies. The sampling of OTUs was found to have the most influence on the tree topologies recovered from the Boyd (2015) dataset, followed by Han *et al.* (2018), and lastly Dieudonné *et al.* (2020), whose topology remained very consistent with DRSCT (original Dieudonné *et al.* [2020] reduced strict consensus tree) when unique taxa were removed (n=27/72; Fig. 4B).

Excluding unique taxa (not found in Boyd [2015]) from Han *et al.* (2018) and Dieudonné *et al.* (2020) did not impact the phylogenetic position of the 'hypsilophodontids' and these taxa remained in a placement consistent with H1 (Hypothesis 1; Fig. 1), within Ornithopoda. In contrast to the results from BSCT (original Boyd [2015] strict consensus tree), removing unique taxa from Boyd (2015) recovers the 'hypsilophodontids' within

Cerapoda, as ornithopods (the H1 position), in both Analysis A.1 (taxa shared with Han *et al.* 2018; Fig. 3) and A.2 (taxa shared with Dieudonné *et al.* 2020; Fig. 4), but still primarily within a monophyletic Thescelosauridae. This suggests that the H2 position of 'hypsilophodontids' found in Boyd (2015) is at least partially related to taxon sampling. Additionally, the original phylogenetic analyses from Boyd (2015) and Han *et al.* (2018) both recovered Heterodontosauridae as an early branching ornithischian clade (in contrast to Dieudonné *et al.* [2020], where the group is recovered within Marginocephalia), but in Analysis A.1 the heterodontosaurids are placed much further up the tree as neornithischians (and even in a sister-group relationship to Marginocephalia in Analysis A.1b). From our results it is clear that OTU choice is capable of significantly altering the position of controversial clades. In particular, it leads us to question the phylogenetic placement of the heterodontosaurids, which seems to be highly dependent on the other taxa included, although it is noteworthy that in analyses of larger-scale dinosaur matrices, heterodontosaurids remain close to the root of Ornithischia (Baron *et al*. 2017b).

While removing unique taxa from the analyses caused numerous changes in topology from the original consensus trees reported, the same is not found when taxa were excluded based on their completeness or instability. Excluding taxa with low completeness scores prior to tree searches (Analysis B) recovered tree topologies very similar to the original studies. The most significant difference is that removing highly incomplete taxa from Han *et al.* (2018) and (to a lesser degree) Dieudonné *et al.* (2020) increased tree resolution. Unstable taxa identified and pruned in the original studies from Han *et al.* (2018) and Dieudonné *et al.* (2020) to increase tree resolution were all found here to be <25% complete, which is likely to explain why *a priori* exclusion of incomplete taxa in Analysis B.2–3 increased the resolution of the strict consensus trees produced. This suggests a relationship between instability and completeness, which is supported by a correlation between completeness values and PCR

(positional congruence reduced) values in taxa from the Han *et al.* (2018) matrix. Within a tree search, a greater proportion of missing data for a taxon (i.e. lower taxon completeness) generates more potential placements for that taxon, which find less consensus, and consequently are typically recovered in a poorly-resolved position (Kearney 2002). This relationship is not found in Boyd (2015), which generally has more incomplete taxa than either H1 matrix (Fig. 5), but fewer unstable taxa (Supplementary material 4, Fig. S23). Instead, this result demonstrates how taxa with relatively few characters scored, can still be highly phylogenetically informative and so should not be excluded from analyses (Kearney & Clark 2003).

Character sampling

The data matrix from Boyd (2015) shares relatively few characters with either Han *et al.* (2018; 30%) or Dieudonné *et al.* (2020; 46%) and it is therefore unsurprising that enigmatic taxa such as 'hypsilophodontids' have been recovered in such conflicting placements within the neornithischian tree. Of those characters that are shared, the majority use different character states and do not capture the same morphological variation. Some of these differences in state partitioning are very subtle, but significantly change how the characters infer evolutionary change and, in some cases, whether these characters are supportive of specific clades.

Character ordering increased the resolution of the strict consensus trees generated from the datasets of Han *et al.* (2018) and Dieudonné *et al.* (2020), without significantly altering the inferred relationships. Ordering multi-state characters where appropriate has been found similarly to increase tree resolution in several previous studies (Slowinski 1993; Grand *et al.* 2013; Rineau 2015).

Analysing craniodental and postcranial character partitions in isolation often recovers different relationships (Mounce *et al.* 2016) and this effect has been shown to be especially marked in non-avian dinosaurs (Li *et al.* 2020). Our analyses demonstrate significant differences between the tree topologies produced from craniodental and postcranial character partitions of the three neornithischian datasets we examined. Across all three matrices, removing either craniodental and postcranial character partitions greatly reduced tree resolution compared to the original consensus trees of Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020). However, in Boyd (2015), postcranial characters were found to yield a more resolved tree topology compared to craniodental characters. This is in contrast to results found for Han *et al.* (2018) and (to a much lesser extent) Dieudonné *et al.* (2020), and also ornithischian phylogenies as a whole (Li *et al.* 2020), in which craniodental characters seem to have more impact on resulting topology. These contrasting results further demonstrate the extent to which the data matrix from Boyd (2015) differs from that of Han *et al.* (2018) and Dieudonné *et al.* (2020).

Character partition exclusion does not affect the position of the 'hypsilophodontids' in Han *et al.* (2018) and Dieudonné *et al.* (2020) and the results of these parsimony analyses remain in agreement with H1. This suggests that characters throughout the skeleton in both H1 matrices have generally congruent phylogenetic signals concerning the position of the 'hypsilophodontids'. Removing dental or hind limb characters from Boyd (2015) recovers the 'hypsilophodontids' within Ornithopoda, in an H1 position and in conflict with the H2 placement originally recovered in Boyd (2015). This would suggest that dental and hind limb partitions both contain characters that are disproportionately more supportive of a H2 topology than characters from other skeletal regions. This was also found to be the case more specifically in individual skeletal elements and characters; excluding only three femoral characters from the Boyd (2015) matrix recovered Thescelosauridae within Cerapoda (an H1

position; Fig. 6), in contrast to results from BSCT. These three femoral characters have completeness scores that are relatively high (range = 29–71%; average = 53%), are not particularly homoplastic (range = 1–3 extra steps), and pertain to distinct morphological features at the proximal (B212, presence/absence of a trench between the greater trochanter and the head of the femur; B213, convex/flat lateral surface of the greater trochanter) and distal end (B223, form of the anterior intercondylar groove) of the femur. B213 is also found to be a synapomorphy of Thescelosauridae, both in this study and in Boyd (2015). In Boyd (2015), this character did not support the clade Jeholosauridae (*Changchunsaurus* + *Haya*, *Jeholosaurus*; Han *et al.* 2012), with *Jeholosaurus* being coded B213:0 (convex lateral surface of the greater trochanter) and *Haya* being coded B213:1 (flat lateral surface of the greater trochanter). This character was later excluded from the modified Boyd (2015) matrix by Madzia *et al.* (2018) and reformulated in Dieudonné *et al.* (2020), where it was found to support Jeholosauridae. These characters, as defined and coded in Boyd (2015), have a strong influence on the topology produced in BSCT and their inclusion significantly alters the inferred interrelationships of ornithischian taxa. The distribution of these characters, the definition of their character states, and the way they are scored should therefore be a focus of future work.

Our results suggest that excluding highly homoplastic or incompletely scored characters does not significantly alter the tree topology or relationships recovered, but reduces the resolution of the consensus trees that are generated. Implied weighting of characters at all strength levels (k =1–5) does not alter the position of 'hypsilophodontids' within Han *et al.* (2018) and Dieudonné *et al.* (2020), although in the latter, the strongest level of weighting (k = 1) does recover Pachycephalosauria (inclusive of Heterodontosauridae, as in DSCT) in a controversial placement, separate from Ceratopsia, at the base of Ornithischia (Supplementary material 4, Fig. S88). The use of implied weighting

in Boyd (2015) results in several major changes compared to the BSCT concerning taxa of interest, most notably recovering: (1) *Hypsilophodon* outside of Cerapoda, rather than as a basal ornithopod; and (2), the elasmarians inside Cerapoda as ornithopods, rather than as a clade of thescelosaurines. These results are more congruent with the updated iteration of the Boyd (2015) matrix by Madzia *et al.* (2018), which also recovers both of these topological features. The Madzia *et al*. (2018) iteration of the Boyd (2015) matrix, along with our own study, demonstrates the labile relationships of these taxa, with numerous changes in relationships possible between Thescelosauridae and Cerapoda, which seem to be particularly marked with elasmarians. The use of implied character weighting is contentious within parsimony analyses (Congreve & Lamsdell 2016; Goloboff *et al.* 2018; Smith 2019b), but it is important to note the effect that homoplastic characters have on these matrices.

Recommendations for future studies

We find that recent analyses that have recovered one of two conflicting placements for the 'hypsilophodontids' are largely distinct in character composition. The independence of these matrices has been beneficial to our understanding of ornithischian phylogenetics, with areas of consensus adding credence to the validity of those common relationships, and areas of disagreement identifying unresolved clades that need more attention (e.g. 'hypsilophodontids', heterodontosaurids). However, we propose that rather than continuing to develop largely independent character matrices that predictably recover either a H1 or H2 placement for 'hypsilophodontids', characters from recent neornithischian data matrices should be combined, critically assessed and updated to reflect current understanding of ornithischian character evolution. Characters should be widely sampled from throughout the skeleton and over-representation of a specific element (e.g. femoral characters in Boyd [2015]) should be limited if possible, although consideration should be given to how

characteristic a particular element is. The characters that support particular clades of interest (e.g. 'Thescelosauridae', Cerapoda and Ornithopoda) or have been shown to be problematic in this study should be assessed to understand how their character construction and state partitioning influence their evolutionary signal.

As the relationships of 'hypsilophodontids' remain relatively fluid with respect to those of 'core' cerapodans, future studies should cast a wide net regarding OTU choice, sampling a broader array of basal forms from the major genasaurian clades and potentially incorporating representatives of other dinosaur clades. Testing whether outgroup choice substantially influences tree topology was beyond the scope of this study, but this is also an important consideration for future studies with several recent major changes to the base of the ornithischian tree (Baron & Barrett 2017; Agnolín & Rozadilla 2018; Dieudonné *et al.* 2020; Müller & Garcia 2020). The phylogenetic analysis from Boyd (2015) was the first to include several poorly known enigmatic neornithischian taxa (e.g. *Atlascopcosaurus, Koreanosaurus, Leaellynasaura, Notohypsilophodon, Qantassaurus*) in a large-scale ornithischian matrix. Several of these taxa (particularly those from the Southern Hemisphere) have not been included in an independent phylogenetic analysis subsequently, possibly due to the incompleteness of some of this material, inadequate published descriptions, their uncertain systematics, or outdated reference material. It is important for future analyses to be as inclusive of these enigmatic taxa as possible, as well as recently described taxa (e.g. Li *et al.* 2019; Yang *et al.* 2020). Importantly, future matrices should include characters representative of the large range of ornithischian taxa included. Characters from Boyd (2015) were predominately (37%; n=120) derived from Scheetz (1999), which focused on the phylogenetic position of *Orodromeus makelai* and did not include any marginocephalian taxa. Because characters derived from these two matrices were not formulated for use on marginocephalian taxa, they may not be appropriate for discerning the phylogenetic positions

of taxa that are relatively labile around the base of Cerapoda, to the extent that they are in Boyd (2015).

Prior to the tree search, safe taxonomic reduction should be utilised to identify taxa that can be excluded without impacting the relationships of the remaining taxa. Here, safe taxonomic reduction identified one taxon to remove from both Boyd (2015; *Thescelosaurus garbanii*) and Han *et al.* (2018; *Yueosaurus tiataiensis*), which in the case of the latter significantly reduced the number of MPTs generated in the tree search. Where it can be justified, characters should be ordered. Although increased resolution does not necessarily increase the accuracy of the resulting topology, it is crucial to discerning phylogenetic relationships (Smith 2019b). Therefore, following the tree search, if the resolution of the strict consensus tree is poor, unstable taxa should be identified and pruned to give a reduced strict consensus (Wilkinson 1994; Pol & Escapa 2009). Taxa should not be removed *a priori* (unless using safe taxonomic reduction), even if they have been identified as being unstable within the topology. Here, removing taxa identified by IterPCR (Pol & Escapa 2009) as unstable *a posterori* (i.e. pruning), resulted in increased tree resolution compared to *a priori* removal. In addition to the recommendations outlined in this study, new computational methods are also being developed to assess conflict between competing topologies derived from morphological datasets (Goloboff & Sereno 2021).

Conclusions

Our study presents the first in-depth assessment and comparison of three recent neornithischian analyses that recover one of two conflicting placements for the ‘hypsilophodontids’: hypothesis 1 (H1), primarily within Cerapoda as a paraphyletic assemblage of basal ornithopods (Han *et al.* 2018; Dieudonné *et al.* 2020); and hypothesis 2 (H2), primarily within the non-cerapodan neornithischian clade, Thescelosauridae (Boyd

2015). Taxon and character manipulation of each data matrix was performed to identify how these factors influenced the recovered topologies.

In all of the analytical permutations applied in this study, the H1 matrices we investigated (Han *et al.* 2018; Dieudonné *et al.* 2020) recovered 'hypsilophodontids' within Cerapoda, in a position consistent with their original H1 placement. By contrast, in the H2 matrix investigated (Boyd 2015), the positions of 'hypsilophodontids' were found to be far more susceptible to the influences of both character and taxon sampling, with several analyses recovering the majority of 'hypsilophodontids' within Cerapoda, in positions more congruent with H1 than H2. Within the Boyd (2015) dataset, characters from femoral and dental regions were found to be disproportionately more supportive of a H2 topology, compared to characters from other regions of the skeleton. Furthermore, removing only three femoral characters from Boyd (2015) resulted in Thescelosauridae being recovered within Ornithopoda. This result emphasises how influential individual characters can be in a data matrix, and how they can have a major impact on our understanding of major evolutionary changes. While the phylogenetic positions of the 'hypsilophodontids' remain contentious and require much further study, our results tend to favour a placement consistent with H1, within Cerapoda.

Acknowledgements

E.E.B. was funded by CENTA (Central England NERC Training Alliance). We thank the Willi Hennig society for sponsorship of the program TNT. Thanks to Paul-Emile Dieudonné for supplying data files used in this study. We would also like to thank the Associate Editor, Jennifer C. Olori (State University of New York Oswego), and two anonymous reviewers for their positive and helpful comments that improved the final manuscript.

References

Agnolín, F. L. & Rozadilla, S. 2018. Phylogenetic reassessment of *Pisanosaurus mertii* Casamiquela, 1967, a basal dinosauriform from the Late Triassic of Argentina. *Journal of Systematic Palaeontology*, **16**, 853–879.

Andrzejewski, K. A., Winkler, D. A. & Jacobs, L. L. 2019. A new basal ornithopod (Dinosauria: Ornithischia) from the Early Cretaceous of Texas. *PLoS ONE*, **14**, e0207935. doi:10.1371/journal.pone.0217232

Bapst, D. & Wagner, P. 2019. 'paleotree' (Paleontological and Phylogenetic Analyses of Evolution), version 3.3.25. https://CRAN.R-project.org/package=paleotree

Baron, M. G. & Barrett, P. M. 2017. A dinosaur missing link? *Chilesaurus* and the early evolution of ornithischian dinosaurs. *Biology Letters*, **13**, 20170220. doi:10.1098/rsbl.2017.0220

Baron, M. G., **Norman, D. B. & Barrett, P. M.** 2017a. Postcranial anatomy of *Lesothosaurus diagnosticus* (Dinosauria: Ornithischia) from the Lower Jurassic of southern Africa: implications for basal ornithischian taxonomy and systematics. *Zoological Journal of the Linnean Society*, **179**, 125–168.

Baron, M. G., **Norman, D. B. & Barrett, P. M.** 2017b. A new hypothesis of dinosaur relationships and early dinosaur evolution. *Nature*, **543**, 501–506.

Barrett, P. M. 2014. Paleobiology of herbivorous dinosaurs. *Annual Review of Earth and Planetary Sciences*, **42**, 207–230.

Barta, D. E. & Norell, M. A. 2021. The osteology of *Haya griva* (Dinosauria: Ornithischia) from the Late Cretaceous of Mongolia. *Bulletin of the American Museum of Natural History*, **445**, 1–112.

Becker, R. A., Wilks, A. R., Brownrigg, R., Minka, T. P. & Deckmyn, A. 2016. maps: Draw Geographical Maps, version 3.3.0. https://CRAN.R-project.org/package=maps

Bell, M. A. & Lloyd, G. T. 2015. strap: an R package for plotting phylogenies against stratigraphy and assessing their stratigraphic congruence. *Palaeontology*, **58**, 379–389.

Bell, P. R., Fanti, F., Hart, L. J., Milan, L. A., Craven, S. J., Brougham, T. & Smith, E. 2019. Revised geology, age, and vertebrate diversity of the dinosaur-bearing Griman Creek formation (Cenomanian), lightning ridge, new south wales, Australia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **514**, 655–671.

Benton, M. J. & Storrs, G. W. 1994. Testing the quality of the fossil record: paleontological knowledge is improving. *Geology*, **22**, 111–114.

Boyd, C. A., Brown, C. M., Scheetz, R. D. & Clarke, J. A. 2009. Taxonomic revision of the basal neornithischian taxa *Thescelosaurus* and *Bugenasaura*. *Journal of Vertebrate Paleontology*, **29**, 758–770.

Boyd, C. A. 2015. The systematic relationships and biogeographic history of ornithischian dinosaurs. *PeerJ*, **3**, e1523. doi:10.7717/peerj.1523

Brazeau, M. D. 2011. Problematic character coding methods in morphology and their effects. *Biological Journal of the Linnean Society*, **104**, 489–498.

Brown, C. M. & Druckenmiller, P. 2011. Basal ornithopod (Dinosauria: Ornithischia) teeth from the Prince Creek Formation (early Maastrichtian) of Alaska. *Canadian Journal of Earth Sciences*, **48**, 1342–1354.

Brown, C. M., Evans, D. C., Ryan, M. J. & Russell, A. P. 2013. New data on the diversity and abundance of small-bodied ornithopods (Dinosauria, Ornithischia) from the Belly River Group (Campanian) of Alberta. *Journal of Vertebrate Paleontology*, **33**, 495–520.

Brusatte, S. L. 2010. Representing supraspecific taxa in higher-level phylogenetic analyses: guidelines for palaeontologists. *Palaeontology*, **53**, 1–9.

Buchholz, P. W. 2002. Phylogeny and biogeography of basal Ornithischia. Pp. 18–34 in D. E. Brown (ed.) *The Mesozoic in Wyoming*. Tate Geological Museum, Casper College, Casper, Wyoming.

Butler, R. J. 2005. The 'fabrosaurid' ornithischian dinosaurs of the upper Elliot Formation (Lower Jurassic) of South Africa and Lesotho. *Zoological Journal of the Linnean Society*, **145**, 175–218.

Butler, R. J. & Upchurch, P. 2007. Highly incomplete taxa and the phylogenetic relationships of the theropod dinosaur *Juravenator starki*. *Journal of Vertebrate Paleontology*, **27**, 253–256.

Butler, R. J., Upchurch, P. & Norman, D. B. 2008. The phylogeny of the ornithischian dinosaurs. *Journal of Systematic Palaeontology*, **6**, 1–40.

Butler, R. J., Jin, L.-Y., Chen, J. & Godefroit, P. 2011. The postcranial osteology and phylogenetic position of the small ornithischian dinosaur *Changchunsaurus parvus* from the Quantou Formation (Cretaceous: Aptian–Cenomanian) of Jilin Province, north-eastern China. *Palaeontology*, **54**, 667–683.

Congreve, C. R. & Lamsdell, J. C. 2016. Implied weighting and its utility in palaeontological datasets: a study using modelled phylogenetic matrices. *Palaeontology*, **59**, 447–462.

Cooper, M. R. 1985. A revision of the ornithischian dinosaur *Kangnasaurus coetzeei* Haughton, with a classification of the Ornithischia. *Annals of the South African Museum*, **95**, 281–317.

Davies, T. W., Bell, M. A., Goswami, A. & Halliday, T. J. 2017. Completeness of the eutherian mammal fossil record and implications for reconstructing mammal evolution through the Cretaceous/Paleogene mass extinction. *Paleobiology*, **43**, 521–536.

Dean, C. D., Mannion, P. D. & Butler, R. J. 2016. Preservational bias controls the fossil record of pterosaurs. *Palaeontology*, **59**, 225–247.

Dieudonné, P.-E., Tortosa T., Torcida Fernández-Baldor, F., Canudo, J. I., Díaz-Martínez, I. 2016. An unexpected early rhabdodontid from Europe (Lower Cretaceous of Salas de los Infantes, Burgos Province, Spain) and a re-examination of basal iguanodontian relationships. *PLoS ONE*, **11**, e0156251. doi:10.1371/journal.pone.0156251

Dieudonné, P. E., Cruzado-Caballero, P., Godefroit, P. & Tortosa, T. 2020. A new phylogeny of cerapodan dinosaurs. *Historical Biology*. doi:10.1080/08912963.2020.1793979

Galton, P. M. 1973. Redescription of the skull and mandible of *Parksosaurus* from the Late Cretaceous with comments on the family Hypsilophodontidae (Ornithischia). *Life Science Contributions, Royal Ontario Museum*, **89**, 1–21.

Goloboff, P. A. & Catalano, S. A. 2016. TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics*, **32**, 221–238.

Goloboff, P. A., Torres, A. & Arias, J. S. 2018. Weighted parsimony outperforms other methods of phylogenetic inference under models appropriate for morphology. *Cladistics*, **34**, 407–437.

Goloboff, P. A. & Sereno, P. C. 2021. Comparative cladistics: identifying the sources for differing phylogenetic results between competing morphology-based datasets. *Journal of Systematic Palaeontology*, **19**, 761–786.

Grand, A., Corvez, A., Duque Velez, L. M. & Laurin, M. 2013. Phylogenetic inference using discrete characters: performance of ordered and unordered parsimony and of three-item statements. *Biological Journal of the Linnean Society*, **110**, 914–930.

Han, F.-L., Barrett, P. M., Butler, R. J. & Xu, X. 2012. Postcranial anatomy of *Jeholosaurus shangyuanensis* (Dinosauria, Ornithischia) from the Lower Cretaceous Yixian Formation of China. *Journal of Vertebrate Paleontology* **32**, 1370–1395.

Han, F.-L., Forster, C. A., Xu, X. & Clark, J. M. 2018. Postcranial anatomy of *Yinlong downsi* (Dinosauria: Ceratopsia) from the Upper Jurassic Shishugou Formation of China and the phylogeny of basal ornithischians. *Journal of Systematic Palaeontology*, **16**, 1159–1187.

Hauser, D. L. & Presch, W. 1991. The effect of ordered characters on phylogenetic reconstruction. *Cladistics*, **7**, 243–265.

Heath, T. A., Hedtke, S. M. & Hillis, D. M. 2008. Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution*, **46**, 239–257.

Herne, M. C., Tait, A. M., Weisbecker, V., Hall, M., Nair, J. P., Cleeland, M. & Salisbury, S. W. 2018. A new small-bodied ornithopod (Dinosauria, Ornithischia) from a deep, high-energy Early Cretaceous river of the Australian–Antarctic rift system. *PeerJ*, **5**, e4113. doi:10.7717/peerj.4113

Herne, M. C., Nair, J. P., Evans, A. R. & Tait, A. M. 2019. New small-bodied ornithopods (Dinosauria, Neornithischia) from the Early Cretaceous Wonthaggi Formation (Strzelecki Group) of the Australian-Antarctic rift system, with revision of *Qantassaurus intrepidus* Rich and Vickers-Rich, 1999. *Journal of Paleontology*, **93**, 543–584.

Huelsenbeck, J. P. 1994. Comparing the stratigraphic record to estimates of phylogeny. *Paleobiology*, **20**, 470–483.

Kearney, M. 2002. Fragmentary taxa, missing data, and ambiguity: mistaken assumptions and conclusions. *Systematic Biology*, **51**, 369–381.

Kearney, M. & Clark, J. M. 2003. Problems due to missing data in phylogenetic analyses including fossils: a critical review. *Journal of Vertebrate Paleontology*, **23**, 263–274.

Li, N., Dai, H., Tan, C., Hu, X., Wei, Z., Lin, Y., Wei, G., Li, D., Meng, L., Hao, B. & You, H. 2019. A neornithischian dinosaur from the Middle Jurassic Xintiangou Formation of Yunyang, Chongqing, China: the earliest record in Asia. *Historical Biology*. doi:10.1080/08912963.2019.1679129

Li, Y., Ruta, M. & Wills, M. A. 2020. Craniodental and postcranial characters of non-avian Dinosauria often imply different trees. *Systematic Biology*, **69**, 638–659.

Lipscomb, D. L. 1992. Parsimony, homology and the analysis of multistate characters. *Cladistics*, **8**, 45–65.

Lloyd, G. T. 2016. Estimating morphological diversity and tempo with discrete character-taxon matrices: implementation, challenges, progress, and future directions. *Biological Journal of the Linnean Society*, **118**, 131–151. doi:10.1111/bij.12746

Maddison, W. P. & Maddison, D. R. 2019. Mesquite: A modular system for evolutionary analysis, version 3.61. http://www.mesquiteproject.org

Madzia, D., Boyd, C. A. & Mazuch, M. 2018. A basal ornithopod dinosaur from the Cenomanian of the Czech Republic. *Journal of Systematic Palaeontology*, **16**, 967–979.

Maidment, S. C. R. 2010. Stegosauria: a historical review of the body fossil record and phylogenetic relationships. *Swiss Journal of Geosciences*, **103**, 199–210.

Maidment, S. C. R. & Barrett, P. M. 2012. Does morphological convergence imply functional similarity? A test using the evolution of quadrupedalism in ornithischian dinosaurs. *Proceedings of Royal Society B*, **279**, 3765–3771.

Mallon, J. C. & Anderson, J. S. 2015. Jaw mechanics and evolutionary paleoecology of the megaherbivorous dinosaurs from the Dinosaur Park Formation (upper Campanian) of Alberta, Canada. *Journal of Vertebrate Paleontology*, **35**, e904323. doi:10.1080/02724634.2014.904323

Mannion, P. D. & Upchurch, P. 2010. Completeness metrics and the quality of the sauropodomorph fossil record through geological and historical time. *Paleobiology*, **36**, 283–302.

Maryanska, T. & Osmólska, H. 1985. On ornithischian phylogeny. *Acta Palaeontologica Polonica*, **30**, 3–4.

Mounce, R. C., Sansom, R. & Wills, M. A. 2016. Sampling diverse characters improves phylogenies: craniodental and postcranial characters of vertebrates often imply different trees. *Evolution*, **70**, 666–686.

Müller, R.T. & Garcia, M. S. 2020. A paraphyletic 'Silesauridae' as an alternative hypothesis for the initial radiation of ornithischian dinosaurs. *Biology Letters*, **16**, 20200417. doi:10.1098/rsbl.2020.0417

Norman, D. B. 1984. A systematic reappraisal of the reptile order Ornithischia. Pp. 157–162 in W.-E. Reif & F. Westphal (eds) *Third symposium on Mesozoic terrestrial ecosystems*. Attempto Verlag, Tubingen.

Norman, D. B., Sues, H.-D., Witmer, L. M. & Coria, R. A. 2004. Basal Ornithopoda. Pp. 393–412 in D. B. Weishampel, P. Dodson & H. Osmólska (eds) *The Dinosauria*. Second Edition. University of California Press, Berkeley.

Paradis, E., Blomberg, S., Bolker, B., Brown, J., Claude, J., Cuong, H. S. & Desper, R. 2019. 'ape' (Analyses of phylogenetics and evolution), version 5.4-1. https://CRAN.R-project.org/package=ape

Pol, D. & Norell, M. A. 2001. Comments on the Manhattan stratigraphic measure. *Cladistics*, **17**, 285–289. doi:10.1006/clad.2001.0166

Pol, D. & Escapa, I. H. 2009. Unstable taxa in cladistic analysis: identification and the assessment of relevant characters. *Cladistics*, **25**, 515–527.

Prevosti, F. J. & Chemisquy, M. A. 2010. The impact of missing data on real morphological phylogenies: influence of the number and distribution of missing entries. *Cladistics*, **26**, 326–339.

Raven, T. J. & Maidment, S. C. 2018. The systematic position of the enigmatic thyreophoran dinosaur *Paranthodon africanus*, and the use of basal exemplifiers in phylogenetic analysis. *PeerJ*, **6**, e4529. doi:10.7717/peerj.4529

Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217–223.

Rineau, V., Grand, A., Zaragüeta, R. & Laurin, M. 2015. Experimental systematics: sensitivity of cladistic methods to polarization and character ordering schemes. *Contributions to Zoology*, **84**, 129–148.

Roberts, E. M., Sampson, S. D., Deino, A. L., Bowring, S. & Buchwaldt, R. 2013. The Kaiparowits Formation: a remarkable record of Late Cretaceous terrestrial environments, ecosystems and evolution in western North America. Pp. 85–106 in A. L. Titus & M. A. Loewen (eds) *At the top of the Grand Staircase: the Late Cretaceous of southern Utah*. Indiana University Press, Bloomington and Indianapolis.

Rosenberg, M. S. & Kumar, S. 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. *Proceedings of the National Academy of Sciences of the USA*, **98**, 10751–10756.

Salgado, L., Canudo, J. I., Garrido, A. C., Moreno-Azanza, M., Martínez, L. C., Coria, R. A. & Gasca, J. M. 2017. A new primitive neornithischian dinosaur from the

Jurassic of Patagonia with gut contents. *Scientific Reports*, **7**, 42778. doi:10.1038/srep42778

Scheetz, R. D. 1999. *Osteology of* Orodromeus makelai *and the phylogeny of basal ornithopod dinosaurs*. PhD Thesis, Montana State University, Bozeman, 186 pp.

Sereno, P. C. 1986. Phylogeny of the bird-hipped dinosaurs. *National Geographic Research*, **2**, 234–256.

Sereno, P. C. 1999. The evolution of dinosaurs. *Science*, **284**, 2137–2147.

Siddall, M. E. 1998. Stratigraphic fit to phylogenies: a proposed solution. *Cladistics*, **14**, 201–208.

Smith, M. R. 2019a. TreeTools: create, modify and analyse phylogenetic trees. doi:10.5281/zenodo.3522725

Smith, M. R. 2019b. Bayesian and parsimony approaches reconstruct informative trees from simulated morphological datasets. *Biology letters*, **15**, 20180632. doi:10.1098/rsbl.2018.0632

Sternberg, C. M. 1937. A classification of *Thescelosaurus*, with a description of a new species. P. 375 in Anon (ed.) *Proceedings of the Geological Society of America for 1936*. Geological Society of America.

Thompson, R. S., Parish, J. C., Maidment, S. C. R. & Barrett, P. M. 2012. Phylogeny of the ankylosaurian dinosaurs (Ornithischia: Thyreophora). *Journal of Systematic Palaeontology*, **10**, 301–312.

Thulborn, R. A. 1971. Origins and evolution of ornithischian dinosaurs. *Nature*, **234**, 75–78.

Tutin, S. L. & Butler, R. J. 2017. The completeness of the fossil record of plesiosaurs, marine reptiles from the Mesozoic. *Acta Palaeontologica Polonica*, **62**, 563–573.

Verrière, A., Brocklehurst, N. & Fröbisch, J. 2016. Assessing the completeness of the fossil record: comparison of different methods applied to parareptilian tetrapods (Vertebrata: Sauropsida). *Paleobiology*, **42**, 680–695.

Walther, M. & Fröbisch, J. 2013. The quality of the fossil record of anomodonts (Synapsida, Therapsida). *Comptes Rendus Palevol*, **12**, 495–504.

Wiens, J. J. 2006. Missing data and the design of phylogenetic analyses. *Journal of Biomedical Informatics*, **39**, 34–42.

Wiens, J. J. & Morrill, M. C. 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. *Systematic Biology*, **60**, 719–731.

Weishampel, D. B. & Heinrich, R. E. 1992. Systematics of Hypsilophodontidae and basal Iguanodontia (Dinosauria: Ornithopoda). *Historical Biology*, **6**, 159–184.

Wilkinson, M. 1992. Ordered versus unordered characters. *Cladistics*, **8**, 375–385.

Wilkinson, M. 1994. Common cladistic information and its consensus representations: Reduced Adams and reduced cladistic consensus trees and profiles. *Systematic Biology*, **43**, 343–368.

Wilkinson, M. 1995. Coping with abundant missing entries in phylogenetic inference using parsimony. *Systematic Biology*, **44**, 501–514.

Wills, M. A. 1999. Congruence between stratigraphy and phylogeny: randomization tests and the gap excess ratio. *Systematic Biology*, **48**, 559–580. doi:10.1080/106351599260148

Xu, X., Forster, C. A., Clark, J. M. & Mo, J. 2006. A basal ceratopsian with transitional features from the Late Jurassic of northwestern China. *Proceedings of the Royal Society B*, **273**, 2135–2140. doi:10.1098/rspb.2006.3566

Yang, Y., Wu, W., Dieudonné, P. E. & Godefroit, P. 2020. A new basal ornithopod dinosaur from the Lower Cretaceous of China. *PeerJ*, **8**, e9832. doi:10.7717/peerj.9832

Zwickl, D. J. & Hillis, D. M. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology*, **51**, 588–598.

Figure captions

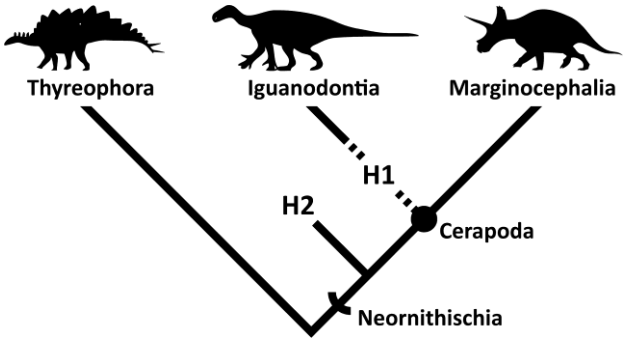

Figure 1. Simplified cladogram showing the two positions where 'hypsilophodontids' have been recovered in recent independent phylogenetic analyses: H1, hypothesis 1 (e.g. Butler *et al.* 2008; Dieudonné *et al.* 2016, 2020; Han *et al.* 2018); H2, hypothesis 2 (e.g. Boyd 2015; Herne *et al.* 2019). Silhouettes from S. Hartman.

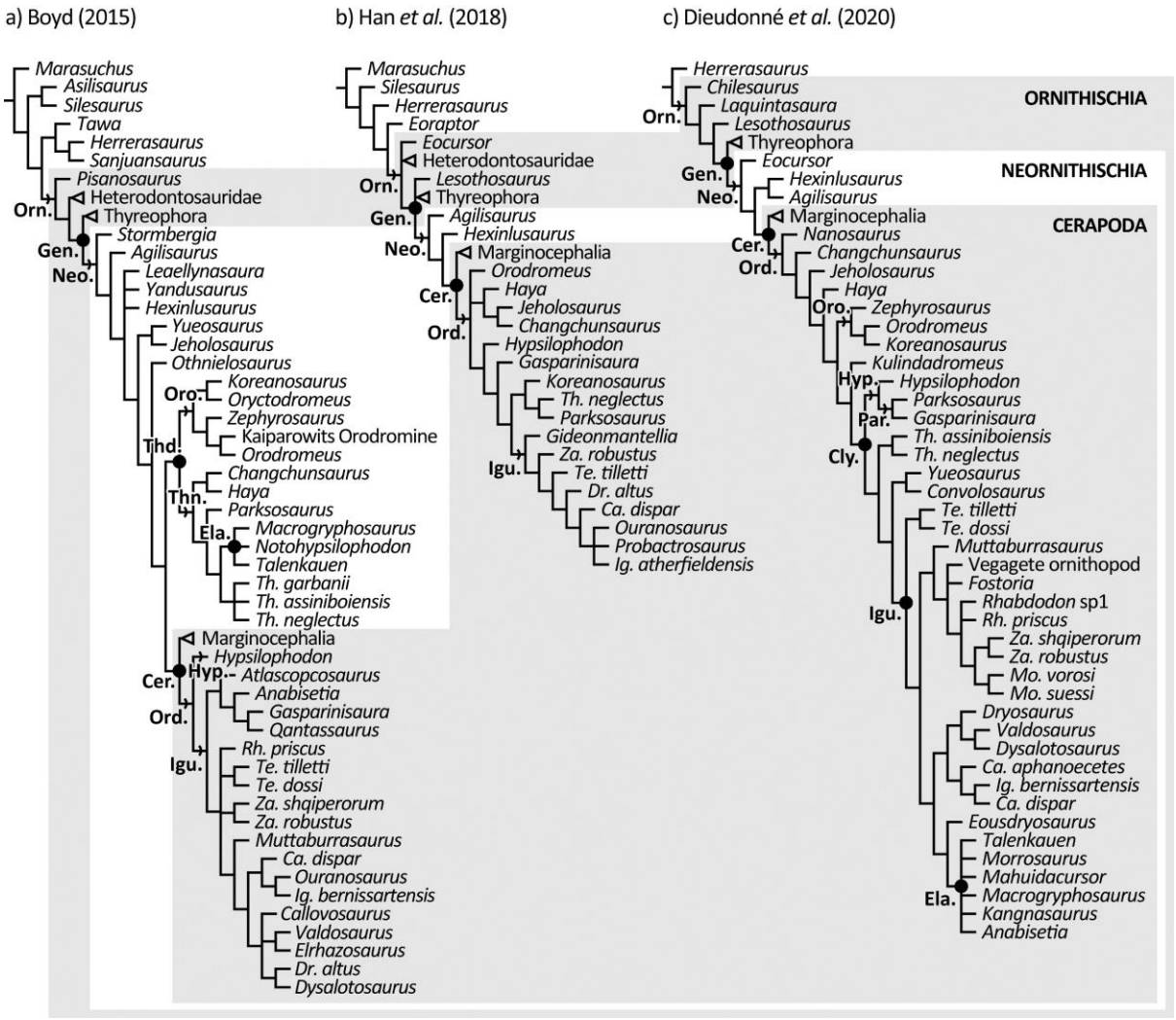

Figure 2. Simplified tree topologies from recent phylogenetic analyses of neornithischian relationships. **A,** strict consensus of the 36 most parsimonious trees (MPTs) generated from the matrix of Boyd (2015); **B**, reduced consensus produced by pruning eight unstable taxa from the 53,376 MPTs generated from the matrix of Han *et al.* (2018); and **C**, reduced consensus produced by pruning *Yandusaurus* from the 176 MPTs generated from the matrix of Dieudonné *et al.* (2020). See Supplementary material 4, Table S1 for clade definitions. **Abbreviations:** ***Ca.,*** *Camptosaurus*; **Cer.**, Cerapoda; **Cly.**, Clypeodonta; ***Dr.***, *Dryosaurus*; **Ela.**, Elasmaria; **Gen.**, Genasauria; **Hyp.**, Hypsilophodontidae; ***Ig.***, *Iguanodon*; **Igu.**, Iguanodontia; ***Mo.***, *Mocholodon*; **Neo.**, Neornithischia; **Ord.**, Ornithopoda; **Orn.**, Ornithischia; **Oro.**, Orodrominae; **Par.**, Parksosauridae; ***Ps.***, *Psittacosaurus*; ***Rh.***, *Rhabdodon*; ***Te.***, *Tenontosaurus*; ***Th.***, *Thescelosaurus*; **Thd.**, Thescelosauridae; **Thn.**, Thescelosaurinae; ***Za.***, *Zalmoxes.*

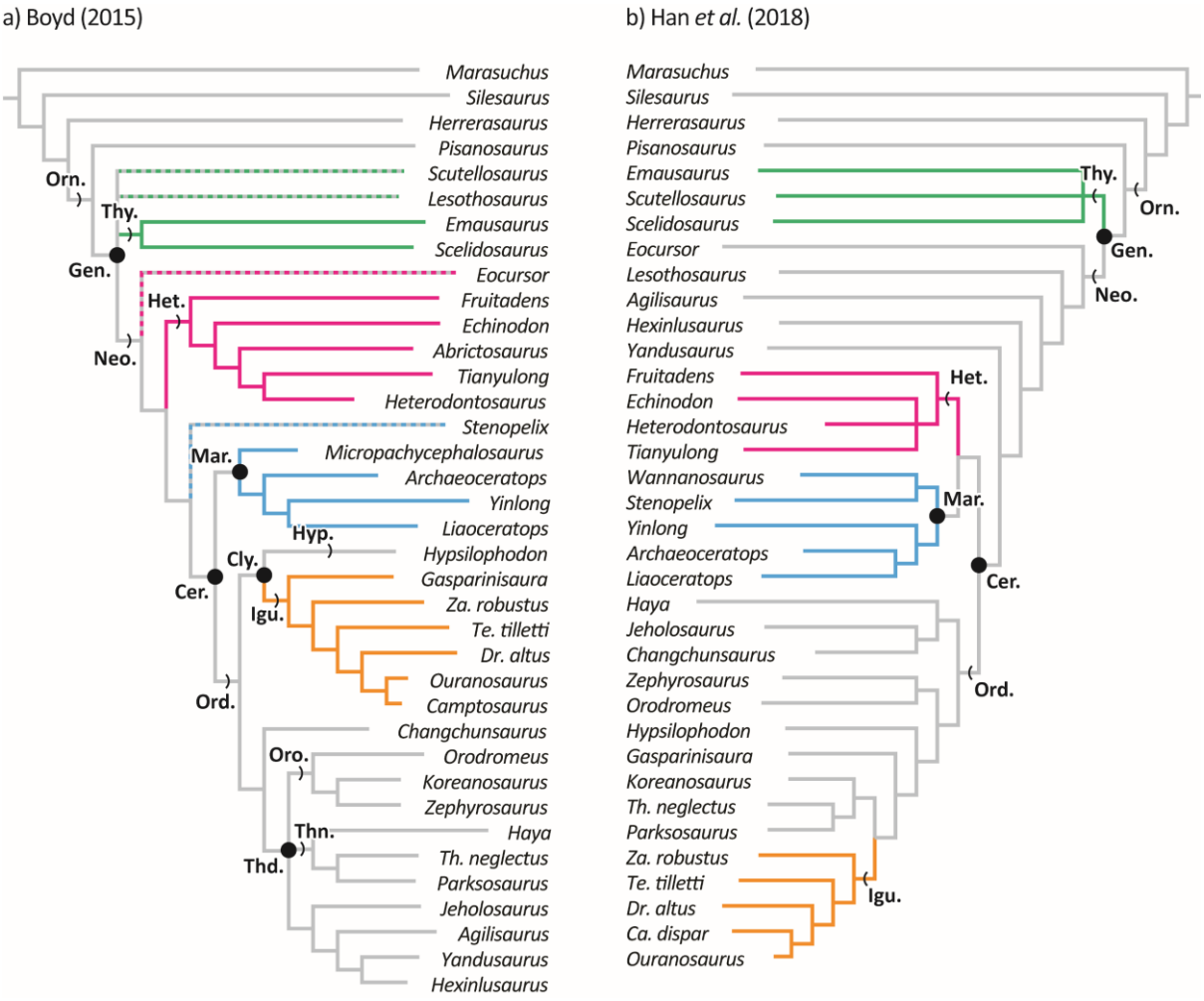

Figure 3. Reduced consensus trees produced using only shared taxa (n=39), following: **A**, pruning of two unstable taxa (*Yueosaurus*, *Wannanosaurus*) from the 66 most parsimonious trees (MPTs) generated from the matrix by Boyd (2015); and **B**, pruning of three unstable taxa (*Abrictosaurus*, *Yueosaurus* and *Micropachycephalosaurus*) from the 164 MPTs generated from the matrix by Han *et al.* (2018). See Supplementary material 4, Table S1 for clade definitions. **Abbreviations:** ***Ca.***, *Camptosaurus*; **Cer.**, Cerapoda; **Cly.**, Clypeodonta; ***Dr.***, *Dryosaurus*; **Gen.**, Genasauria; **Het.**, Heterodontosauridae (pink); **Hyp.**, Hypsilophodontidae; **Igu.**, Iguanodontia (yellow); **Mar.**, Marginocephalia (blue); **Neo.**, Neornithischia; **Ord.**, Ornithopoda; **Orn.**, Ornithischia; **Oro.**, Orodrominae; ***Te.***, *Tenontosaurus*; ***Th.***, *Thescelosaurus*; **Thd.**, Thescelosauridae; **Thn.**, Thescelosaurinae; **Thy.**,

Thyreophora (green); ***Za***., *Zalmoxes*. Dashed lines indicate the group in which a taxon was recovered in the original analyses of Boyd (2015) and Han *et al.* (2018), respectively.

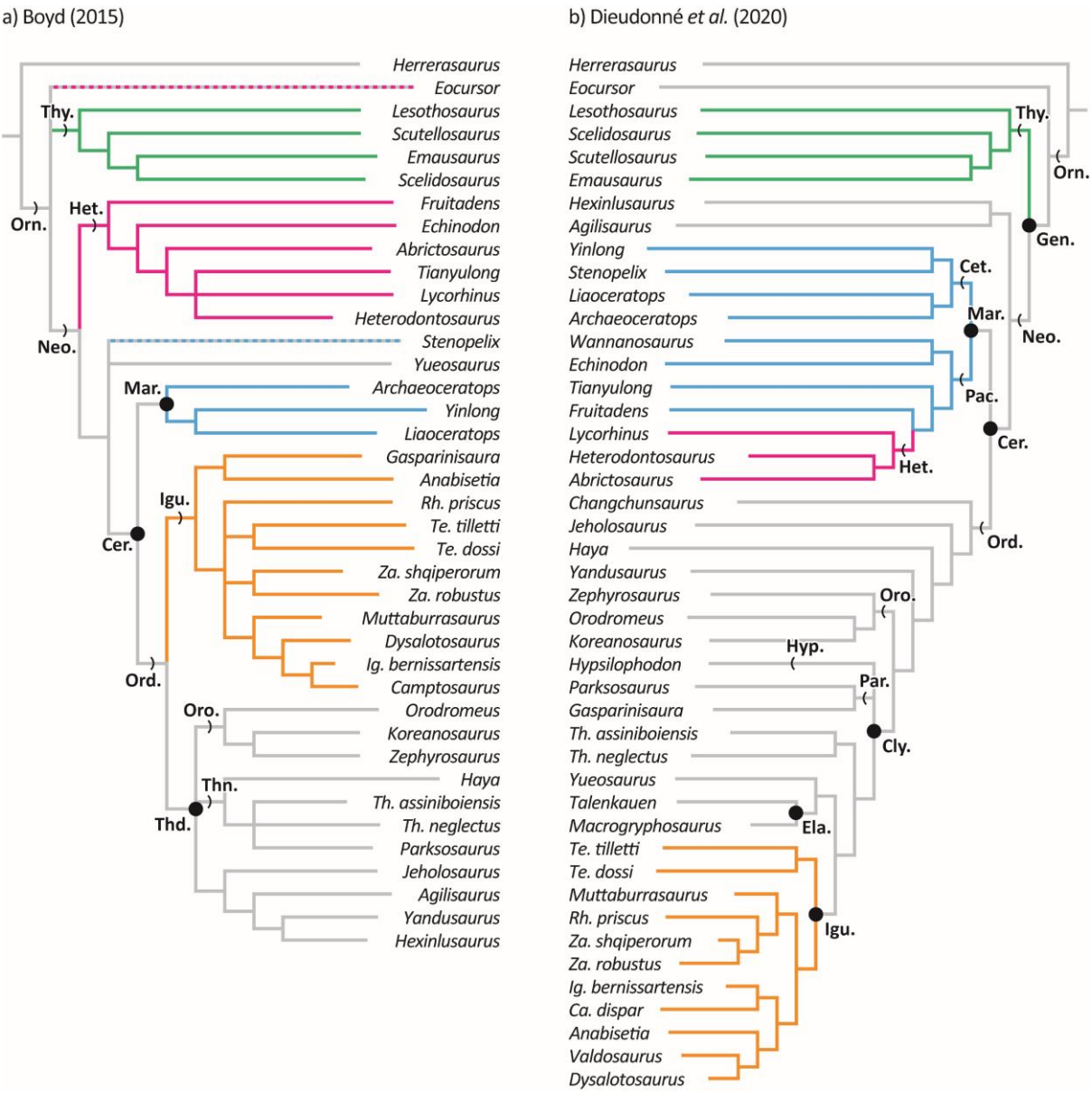

Figure 4. Consensus trees produced using only shared taxa (n=45): (a) following pruning of six unstable taxa (*Changchunsaurus*, *Hypsilophodon*, *Macrogryphosaurus*, *Talenkauen*, *Valdosaurus*, *Wannanosaurus*) from the 180 most parsimonious trees (MPTs) generated from the matrix of Boyd (2015); and, (b) the strict consensus of the two MPTs generated from the matrix of Dieudonné *et al*. (2020). **Abbreviations: *Ca.***, *Camptosaurus*; **Cer.**, Cerapoda;

Cly., Clypeodonta; **Ela.**, Elasmaria; **Gen.**, Genasauria; **Het.**, Heterodontosauridae (pink); **Hyp.**, Hypsilophodontidae; **Igu.**, Iguanodontia (orange); ***Ig.***, *Iguanodon*; **Mar.**, Marginocephalia (blue); **Neo.**, Neornithischia; **Ord.**, Ornithopoda; **Orn.**, Ornithischia; **Oro.**, Orodrominae; **Par.**, Parksosauridae; ***Rh.***, *Rhabdodon*; ***Te.***, *Tenontosaurus*; ***Th.***, *Thescelosaurus*; **Thd.**, Thescelosauridae; **Thn.**, Thescelosaurinae; **Thy.**, Thyreophora (green); ***Za.***, *Zalmoxes*. Dashed lines indicate the group in which a taxon was recovered in the original analyses of Boyd (2015) and Dieudonné *et al.* (2020), respectively.

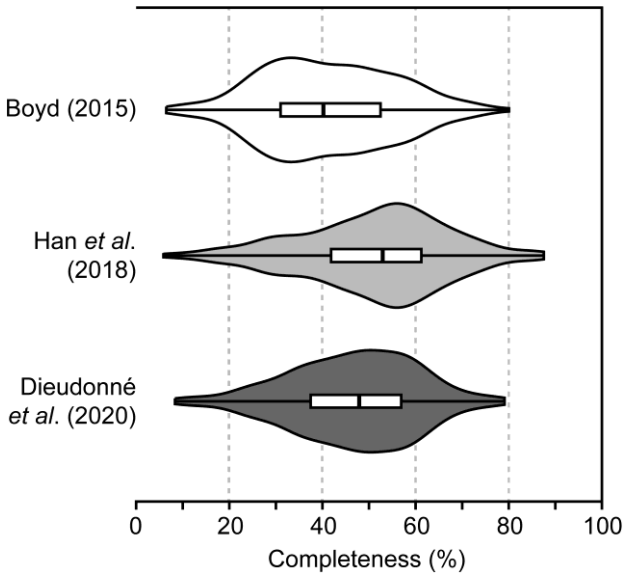

Figure 5. Distribution of completeness scores for scored taxa in the data matrices of Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020).

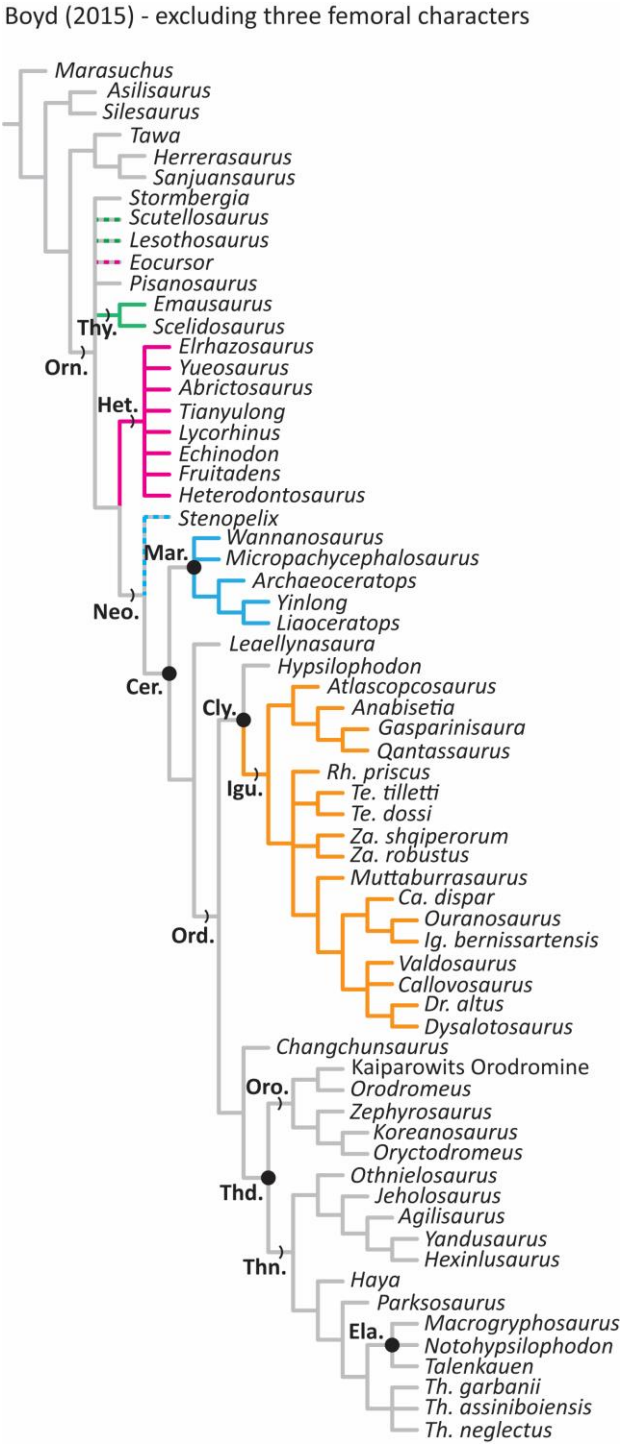

Figure 6. Strict consensus tree of 600 most parsimonious trees generated from Boyd (2015) when femoral characters #212, #213 and #223 are removed. **Abbreviations: Cer.**, Cerapoda; **Cly.**, Clypeodonta; **Ela.**, Elasmaria; **Hyp.**, Hypsilophodontidae; ***Ig.***, *Iguanodon*; **Igu.**, Iguanodontia (orange); **Mar.**, Marginocephalia (blue); **Neo.**, Neornithischia; **Ord.**, Ornithopoda; **Orn.**, Ornithischia; **Oro.**, Orodrominae; ***Te.***, *Tenontosaurus*; ***Th.***,

Thescelosaurus; **Thd.**, Thescelosauridae; **Thn.**, Thescelosaurinae; **Thy.**, Thyreophora (green); ***Za.***, *Zalmoxes*. Coloured dashed lines indicate the group an OTU was recovered in the original analysis of Boyd (2015). Dashed lines indicate the group in which a taxon was recovered in the original analysis of Boyd (2015).

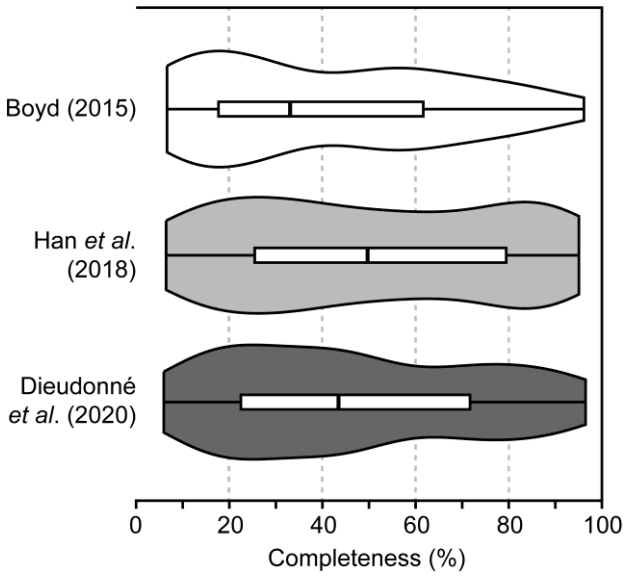

Figure 7. Distribution of completeness scores for individual characters in data matrices from Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020).

Table 1. Summary of the three data matrices analysed in this study and methods originally used to analyse them. Abbreviations: H1, hypothesis 1; H2, hypothesis 2; MPTs, most parsimonious trees; OTU, operational taxonomic unit; TBR, tree bisection reconnection.

	Boyd (2015)	Han *et al.* (2018)	Dieudonné *et al.* (2020)
No. of OTUs	65	72	72
No. of characters	255	380	342
Character weighting	Equal	Equal	Equal
No. of ordered characters	0	21	4
Tree search methods	Traditional tree search using TBR swapping algorithm with 10,000 replications, holding 10,000 trees per replication.	Traditional tree search using TBR swapping algorithm with 1,000 replications, holding 100 trees per replication. Subsequent TBR swapping of trees in RAM.	Traditional tree search using TBR swapping algorithm with 1,000 replications, holding 10 trees per replication. Subsequent TBR swapping of trees in RAM.
MPTs generated	36	53,376	176
Phylogenetic position of the 'hypsilophodontids'	H2	H1	H1

Table 2. Summary of phylogenetic analyses conducted in this study. 'X' denotes the data matrix being analysed, either: B2015, Boyd (2015); H2018, Han *et al*. (2018); or, D2020, Dieudonné *et al*. (2020).

Analysis Name	Description
Analysis A.1	Excluding OTUs from Boyd (2015) and Han *et al.* (2018) that are not common in both matrices
Analysis A.2	Excluding OTUs from Boyd (2015) and Dieudonné *et al.* (2018) that are not common in both matrices
Analysis B.X.1	Excluding OTUs under 10% complete from X
Analysis B.X.2	Excluding OTUs under 20% complete from X
Analysis B.X.3	Excluding OTUs under 30% complete from X
Analysis C.X	Safe taxonomic reduction of OTUs from X
Analysis D.X	Removal of unstable OTUs from X
Analysis E.X	Unordering all characters from X
Analysis F.X.1	Excluding craniodental characters from X
Analysis F.X.2	Excluding postcranial characters from X
Analysis G.X.1	Excluding rostral characters from X
Analysis G.X.2	Excluding palatal characters from X
Analysis G.X.3	Excluding cranial roof characters from X
Analysis G.X.4	Excluding skull base characters from X
Analysis G.X.5	Excluding braincase characters from X
Analysis G.X.6	Excluding mandible characters from X
Analysis G.X.7	Excluding dentition characters from X
Analysis G.X.8	Excluding vertebral column & rib characters from X
Analysis G.X.9	Excluding pectoral girdle & sternum characters from X
Analysis G.X.10	Excluding forelimb characters from X
Analysis G.X.11	Excluding pelvic girdle characters from X
Analysis G.X.12	Excluding hind limb characters from X
Analysis H.X	Excluding the most character-rich skeletal element from X
Analysis I.X.1	Implied character weighting k=1 of X
Analysis I.X.2	Implied character weighting k=3 of X
Analysis I.X.3	Implied character weighting k=5 of X
Analysis J.X.1	Excluding characters homoplasy > 95th percentile from X
Analysis J.X.2	Excluding characters homoplasy > 85th percentile from X
Analysis J.X.3	Excluding characters homoplasy > 75th percentile from X
Analysis K.X.1	Excluding characters under 10% complete from X
Analysis K.X.2	Excluding characters under 20% complete from X
Analysis K.X.3	Excluding characters under 30% complete from X

Table 3. Results of stratigraphic congruence tests of maximum parsimony trees (MPTs) from re-analysis of matrices by Boyd (2015), Han *et al*. (2018) and Dieudonné *et al.* (2020).

	SCI	GER	MSM*	MIG
Boyd (2015; 36 MPTs)				
Median	0.5317	0.8298	0.1344	1337.52
Range	0.5238–0.5397	0.8160–0.8304	0.1256–0.1349	1333.32–1431.52
Han *et al*. (2018; 53,367 MPTs)				
Median	0.5143	0.8352	0.1291	1316.62
Range	0.4714–0.5714	0.8094–0.8484	0.1136–0.1388	1224.87–1496.12
Dieudonné *et al*. (2020; 176 MPTs)				
Median	0.5	0.8201	0.1121	1429.33
Range	0.4714 – 0.5286	0.8181 – 0.8248	0.1109 – 0.1147	1396.43-1444.03

Table 4. Results of stratigraphic congruence tests of strict consensus trees (SCT) and reduced censuses trees (RCT) from re-analysis of matrices by Boyd (2015), Han *et al.* (2018) and Dieudonné *et al.* (2020).

	SCI	GER	MSM*	MIG
Boyd (2015)				
SCT	0.5317	0.8298	0.1344	1337.52
Han *et al*. (2018)				
SCT	0.6364	0.7531	0.0901	1887.42
RCT	0.5455	0.8266	0.1365	1245.73
Dieudonné *et al*. (2020)				
SCT	0.5455	0.7917	0.0983	1630.03
RCT	0.5082	0.8177	0.1116	1436.03

Table 5. Results of Pearson's r correlation test between logged character completeness and logged positional congruence (reduced; PCR) of taxa from Boyd (2015), Han *et al.* (2018) and Dieudonné *et al*. (2020). Significant *p*-values are italicised.

	Pearson's r	p-value
Boyd (2015)	-0.0647	0.6084
Han *et al*. (2018)	0.3168	*0.0067*
Han *et al.* (2018; minus outliers)	0.1801	0.1512
Dieudonné *et al*. (2020)	0.0975	0.4153

Table 6. Proportion of characters relating to a specific skeletal region relative to all characters, from the matrices by Boyd (2015; n=255), Han *et al.* (2018; n=380) and Dieudonné *et al*. (2020; n=342).

	Relative proportion of total characters (%)		
	Boyd (2015)	Han *et al*. (2018)	Dieudonné *et al*. (2020)
Craniodental vs. Post-cranial			
Craniodental	55.69	59.47	56.43
Post-cranial	44.31	40.53	43.57
Skeletal region			
Rostrum	10.59	10.53	10.53
Palate	0	2.63	1.46
Cranial roof	4.71	7.89	6.43
Lateral skull base	14.12	12.11	11.11
Braincase	4.71	4.21	6.14
Mandible	8.24	11.05	8.77
Dentition	12.16	8.42	10.53
Vertebral column	6.67	8.42	7.60
Pectoral girdle & sternum	3.14	2.89	3.51
Forelimbs	6.67	6.84	7.02
Pelvic girdle	10.20	11.32	10.82
Hind limbs	16.47	8.68	12.87
N/A	2.35	5.00	3.22