

# Chapter 3

## The Impact of Unstable Taxa in Coelurosaurian Phylogeny and Resampling Support Measures for Parsimony Analyses

DIEGO POL<sup>1</sup> AND PABLO A. GOLOBOFF<sup>2</sup>

### ABSTRACT

Paleontological datasets often have large amounts of missing entries that result in multiple most parsimonious trees. Highly incomplete and conflictive taxa produce a collapsed strict consensus and several methods have been developed for identifying these unstable or rogue taxa in optimal trees derived from phylogenetic analyses. In addition to decreasing consensus resolution, incomplete or conflictive taxa can also severely affect the support values of phylogenetic analysis in paleontological datasets. Here, we explore a protocol for the identification of taxa that decrease jackknife support values in parsimony analysis. The taxa identified are excluded from majority rule jackknife trees, revealing nodes that have either low or high support irrespective of the uncertainties in the placement of unstable taxa. A recently published dataset of coelurosaurian relationships based on 164 taxa and 853 characters is explored using this protocol; our protocol detects a total of 40 unstable taxa as the most detrimental for node supports. Major clades that are well supported in the reduced jackknife tree include Coelurosauria, Maniraptoriformes, Compsognathidae, Ornithomimosauria, Alvarezsauroidea, Therizinosauria, Oviraptorosauria. Clades with moderate support instead include Maniraptora, Pennaraptora, Paraves, Dromaeosauridae, Troodontidae, Anchiornithinae, and early-diverging clades of Avialae.

### INTRODUCTION

Morphological datasets that include a large number of extinct taxa are usually characterized by copious amounts of missing entries. The abundance of missing data in these datasets has been regarded as problematic for phylogenetic analyses, since the early days of quantitative cladistics (Gauthier, 1986; Wilkinson and Benton, 1995). The presence of taxa with abundant missing entries has been linked to searches that find multiple optimal trees in parsimony analyses and the related computational difficulties of dealing

with thousands of trees (which were very problematic for early phylogenetic software). A subsequent problem in these cases is how the multiple optimal trees can be efficiently summarized given that the strict consensus is usually highly collapsed due to the alternative positions of wildcard or rogue taxa. The role of reduced consensus methods (Wilkinson, 1994) has become increasingly important in recent years, and several methods have been proposed and implemented for detecting rogue taxa in a collection of optimal trees (Goloboff et al., 2008; Pol and Escapa, 2009; Goloboff and Szumik, 2015).

---

<sup>1</sup> CONICET, Museo Paleontológico Egidio Feruglio, Trelew, Argentina.

<sup>2</sup> CONICET, Unidad Ejecutora Lillo (UEL), S.M. Tucumán, Argentina.

A second level of problems introduced by the presence of copious missing entries is related to their influence in support values. It is commonly the case that highly incomplete taxa can be more easily placed in alternative positions than more complete taxa. This obviously affects not only the support with which a taxon is placed in the phylogenetic tree but also the support values of many adjacent nodes (Wilkinson, 1996; Wilkinson et al., 2000). The empirical outcome of this is that paleontological datasets are also characterized by the presence of low support values for most of the nodes recovered in the consensus tree. Some methods exist for assessing the role of rogue taxa for parsimony measures such as Bremer or decay support (e.g., double decay; Wilkinson et al., 2000). Also, alternative resampling methods that do not eliminate characters and thus produce lower estimates of support only in the presence of actual character conflict have been explored (e.g. the “nozeroweight” option for resampling in TNT, used by Pei et al., in press). Recently, however, most efforts have been focused on the development of ways to detect wildcard taxa that decrease bootstrap support (Pattengale et al., 2011; Aberer and Stamatakis, 2011; Aberer et al., 2013).

Here, we explore the application of a more detailed protocol that identifies unstable taxa that decrease support measures, based on resampling procedures (i.e., jackknife or bootstrap) implemented with a script for TNT (Goloboff et al., 2008), that combines several of the options for identifying rogue taxa that already exist in that program. We employ this procedure for a comprehensive phylogenetic analysis of Coelurosauria, using the Theropod Working Group (TwiG) matrix published by Pei et al. (in press). This dataset has an extensive taxon sampling (164 taxa), and therefore provides an ideal case for testing the impact of fragmentary taxa on support measures. In particular, the TwiG dataset includes a dense sampling of pennaraptoran coelurosaurians that have been the focus of recent systematic debates, including the interrelationships of some of its major clades as well as

the affinities of some small but conflictive groups such as scansoriopterygids (Xu et al., 2011, 2015, 2017; Turner et al., 2012; Agnolín and Novas, 2013; O'Connor and Sullivan, 2014) and unenlagiines (Turner et al., 2007; Turner et al., 2012; Agnolín and Novas, 2013; Brusatte et al., 2014). The application of the new protocol allows analyzing the varying degrees of clade support within this group and distinguishing between low support caused by fragmentary taxa, and low support due to underlying character conflict and/or lack of sufficient phylogenetic data.

## MATERIALS AND METHODS

### PHYLOGENETIC ANALYSIS

The phylogenetic dataset (Pei et al., in press) has 164 taxa scored across 853 characters and its parsimony analysis is best carried out using the “New Technology Searches” option in TNT (Goloboff et al., 2008). This strategy was applied in a first phase until 50 hits to minimum length were achieved (command: *xmult = hits 50*), resulting in trees of 3424 steps. The application of traditional heuristic searches (multiple replicates of Wagner trees followed by TBR branch swapping) is possible, but then finding optimal trees requires longer search times than new technology searches. The strict consensus of this analysis is well resolved, with two relatively large polytomies at the base of Avialae and in Dromaeosauridae (fig. 1). These polytomies are caused by four unstable taxa that take multiple positions among the MPTs (i.e., *Yurgovuchia*, *Acheroraptor*, *Velociraptor osmolskae*, and *Archaeopteryx* Haarlem). Ten other taxa are also identified as unstable by IterPCR (Pol and Escapa, 2009) applied to the MPTs, as implemented in TNT (see Goloboff and Szumik, 2015). Ignoring all the unstable taxa results in a well-resolved, strict reduced-consensus tree (see resolved nodes in fig. 1).

It is important to note that the unstable taxa are pruned from the trees, but they are *not* eliminated from the matrix or the tree searches at any point. The elimination of taxa from the trees

amounts to representing those parts of the results that are more useful, while eliminating taxa from the matrix amounts to ignoring the evidence (in the form of character combinations) provided by those taxa (see discussion in Goloboff and Szumik, 2015: 100–101, and fig. 6).

#### RESAMPLING SUPPORT MEASURES

In this paper we compare jackknife support values obtained using the default settings in TNT (including all taxa and using group frequencies on the majority rule consensus) with jackknife frequencies obtained with the same procedure but on a reduced majority rule consensus tree (ignoring the alternative position of unstable taxa).

We opted for this comparison for the sake of simplicity to highlight the effect of unstable taxa on resampling support measures. However, we note that there are alternative ways for summarizing resampling measures, such as GC frequencies (rather than raw frequencies) and/or measuring frequencies for the nodes present in the strict consensus of the MPTs (rather than those appearing in the majority rule consensus of the resampling procedure).

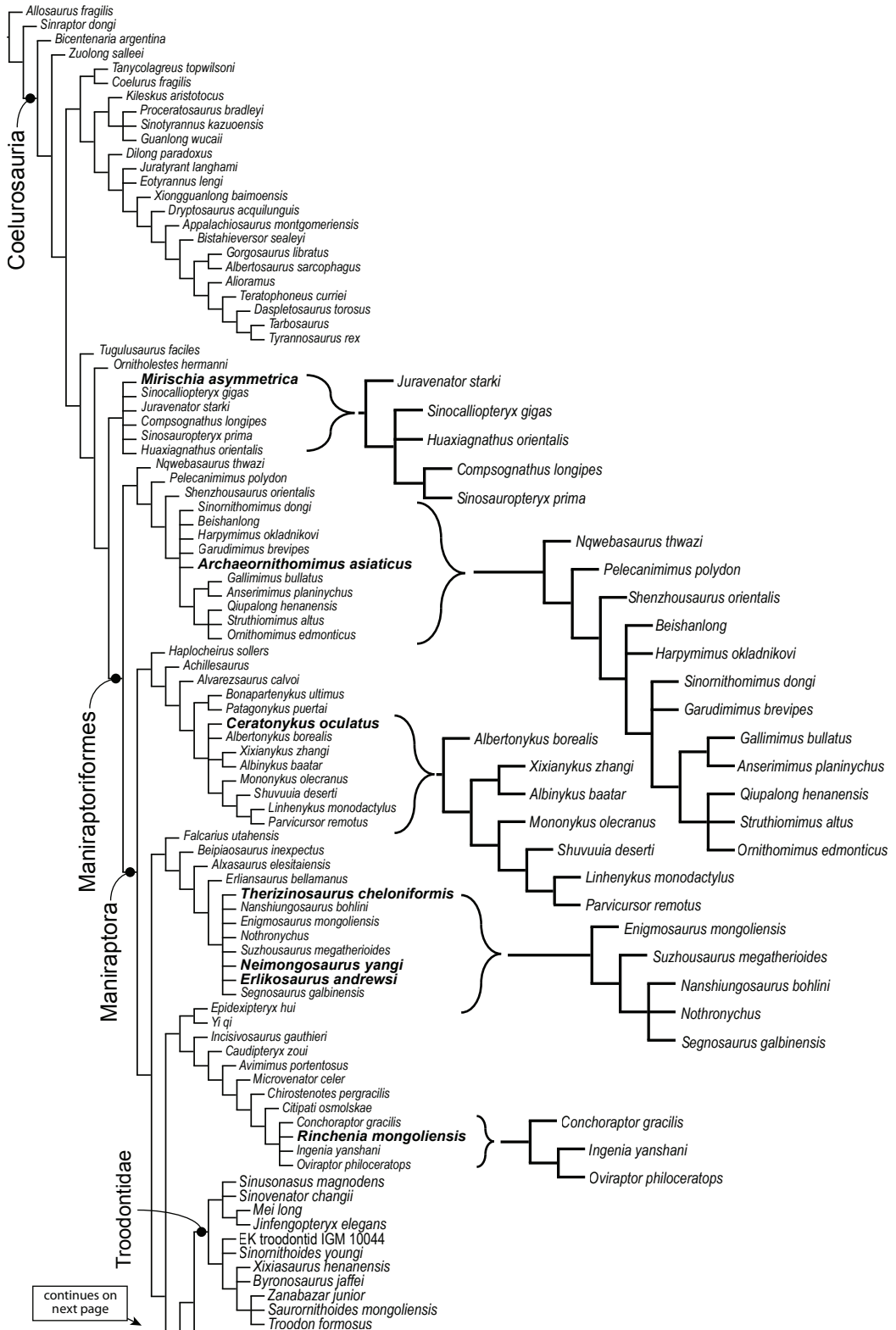
The frequencies of the majority rule consensus obtained after performing a default jackknife analysis on the dataset of Pei et al. (in press) are low for most nodes. A raw frequency majority rule tree of the jackknife analysis is highly collapsed and has only 68 nodes, only 33 of which have support values above 75% (fig. 2). The jackknife majority rule tree shows a large polytomy at the base of Pennaraptora that involves the relationships of 38 terminal taxa and six moderately supported clades. Two important clades of Paraves (Dromaeosauridae and Avialae) are collapsed into the large early-diverging polytomy of Pennaraptora. The frequencies of the groups resolved in the reduced consensus (fig. 1) can be calculated for groups with frequency below 50%, but even in this case, the support values of Dromaeosauridae and Avialae are very low when all the remaining taxa are included in the summary tree. This situation, in which many clades

of interest are collapsed in the majority rule tree derived from the resampling replicates, is in fact a common result in paleontological datasets (in particular when they are constructed using an extensive taxon-sampling regime).

#### IDENTIFYING UNSTABLE TAXA FOR RESAMPLING SUPPORT ANALYSIS

A resampling procedure (e.g., jackknife or bootstrap) normally involves conducting at least 100 pseudoreplicates. In each of these replicates the characters of the original matrix are resampled at random so that a modified (perturbed) matrix is obtained and a tree search is conducted on this modified matrix (fig. 3). The difference between alternative resampling support measures is simply how the resampling of characters is performed (e.g., bootstrap, jackknife, symmetric; Farris et al., 1996; Goloboff et al., 2003). Regardless of how this is done, the tree search conducted for each pseudoreplicate results in a set of most parsimonious trees and, therefore, after finishing the 100 pseudoreplicates, there are 100 sets of most parsimonious trees.

As noted previously (Goloboff et al., 2003; Simmons and Freudenstein, 2011), some phylogenetic software (e.g., PAUP) calculates the bootstrap/jackknife frequencies by weighting groups according to their frequency within each pseudoreplicate, which can easily produce inappropriate estimates of support. This problem is prevented if the results for each individual pseudoreplicate are instead summarized by means of a strict consensus. Thus, TNT automatically calculates the strict consensus of the optimal trees for each pseudoreplicate, subsequently deriving the bootstrap/jackknife frequencies from the majority rule consensus tree of the 100 strict consensuses (each representing one pseudoreplicate; see fig. 3). Problems similar to those resulting from weighting groups according to their frequency within pseudoreplicates occur if a single tree is found or saved during the search for each pseudoreplicate (Goloboff and Pol 2005), with the possibility of bias in tree searches



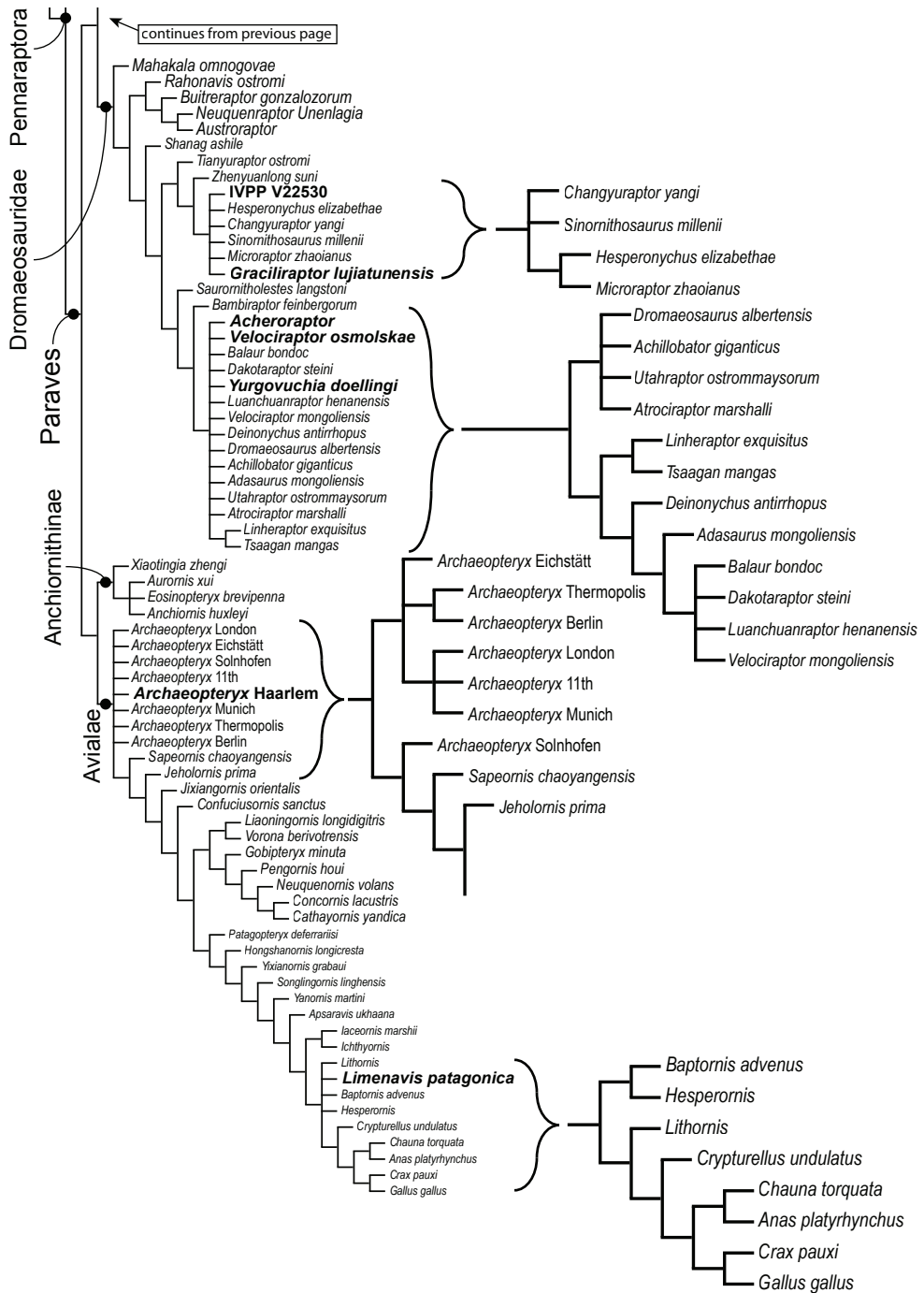


FIG. 1. Strict consensus for the dataset of Pei et al., in press. Details of nodes resolved in the reduced consensus tree are shown, after excluding the 14 taxa that are unstable among the MPTs by the IterPCR implementation of TNT (Goloboff and Szumik, 2015). Taxa in bold in the strict consensus are the unstable taxa detected by IterPCR.

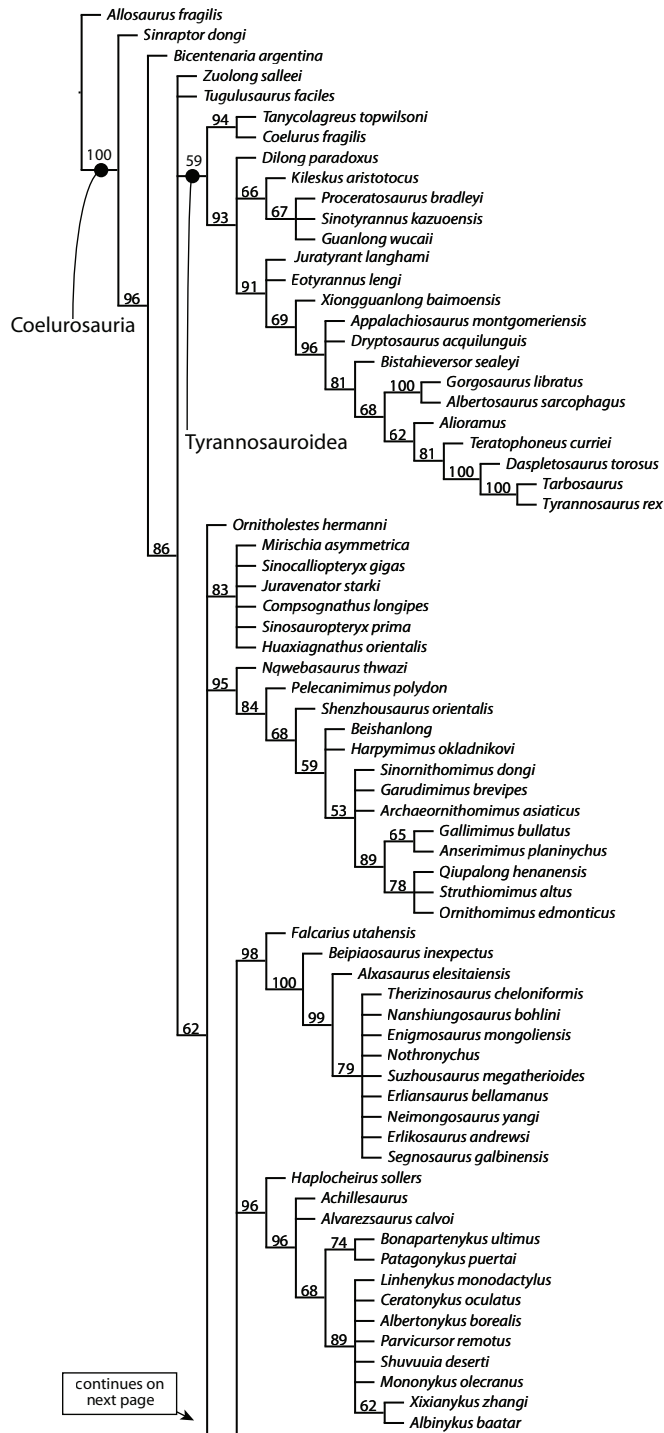
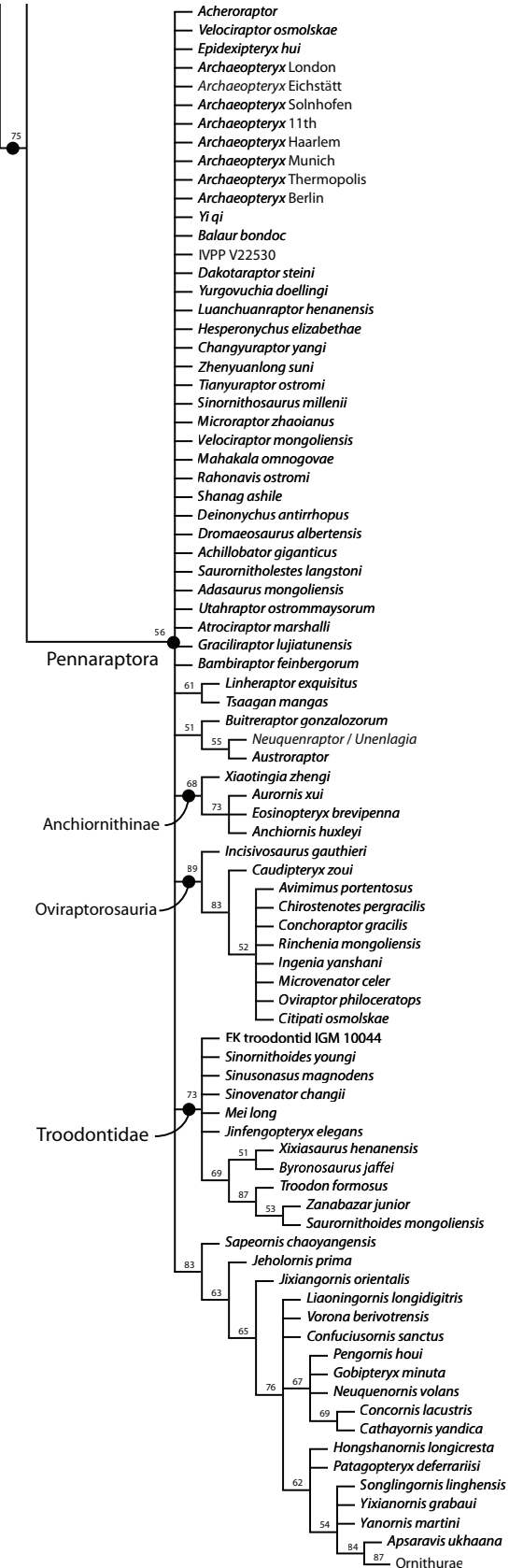


FIG. 2. Absolute frequency jackknife tree for the dataset of Pei et al. (in press) including all taxa as obtained in TNT.

continues from previous page

Maniraptora



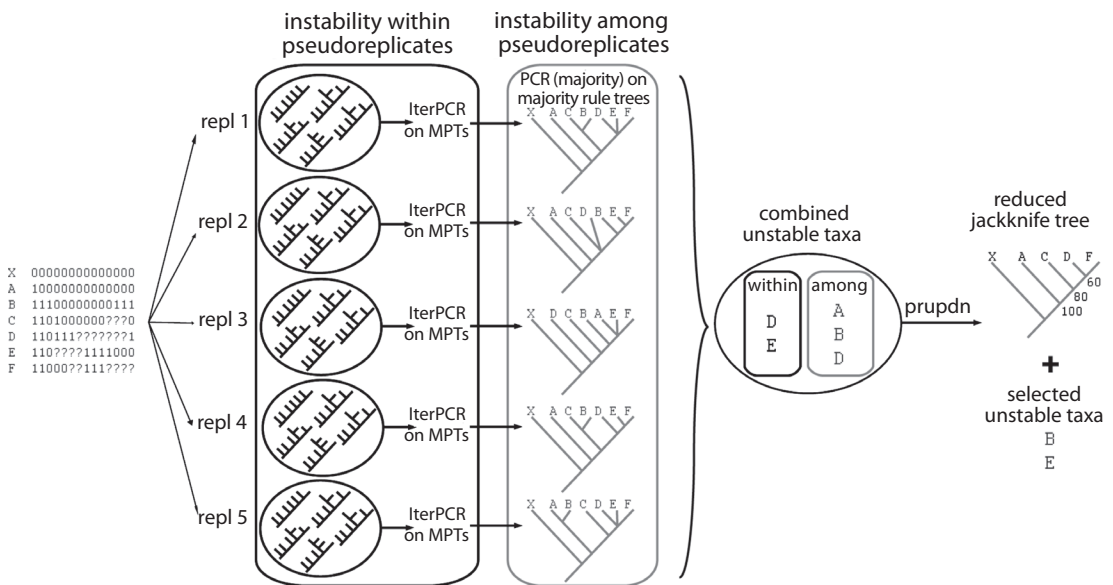


FIG. 3. Resampling support protocol indicating the steps in which the analysis of unstable taxa is applied, as implemented in the TNT script *pcrjak.run* (see <http://www.lillo.org.ar/phylogeny/tnt/scripts/pcrjak.run>).

(Goloboff and Simmons 2014) exacerbating the problem. The best assessments of support, therefore, are obtained when multiple trees are found for each pseudoreplicate, and they are summarized by a strict consensus tree. The multiple trees for each pseudoreplicate can be found during resampling either explicitly (by actual searches) or implicitly (i.e., finding a single tree approaching optimality, then collapsing groups lost on TBR rearrangements producing equally optimal trees; see Goloboff et al., 2008).

The identification of taxa that affect resampling support measures can be approached in several ways. Some authors (Wilkinson, 1996; Pattengale et al., 2011; Aberer and Stamatakis, 2011; Aberer et al., 2013) have analyzed a single combined set of topologies, derived from all the pseudoreplicates. This is made more difficult when the results for each of the individual pseudoreplicates are summarized by means of a strict consensus: the unstable taxa cannot be identified a posteriori of the calculation of the strict consensus. To illustrate the problem, consider a dataset with a terminal with a missing entry in every single character. If the strict consensus is

produced from the multiple trees for each pseudoreplicate, every one of those will be a complete bush. Subsequent pruning of terminals from those trees cannot produce an improved resolution; rather, the terminals must be pruned from the original tree(s) for each pseudoreplicate, not from their (full) strict consensuses.

In the original analysis of Pei et al. (in press), the strategy used to cope with the problem just discussed had been: (1) using a fixed sequence of random seeds to produce the pseudoreplicate datasets and saving a single binary tree per pseudoreplicate; (2) using approximate methods to identify the effect on the majority rule of different taxon prunings in the resulting set of binary trees; and (3) once a set of unstable taxa is identified, the pseudoreplicate datasets are generated again with the same sequence of random seeds and the binary trees are collapsed by applying the implicit TBR-collapsing described above on the resampled dataset. This avoids having to search again for a large number of pseudoreplicate datasets; the optimal trees will not change and they can simply be used to collapse (via TBR) those groups that are unsupported by the pseudoreplicate.



## A COMBINED METHOD FOR DETECTING INSTABILITY

The present paper explores an alternative way to deal with the problem that unstable taxa cannot be identified from the strict consensus resulting from the multiple trees produced by every pseudoreplicate, taking advantage of the fact that the scripting facilities offered by TNT make it possible to easily explore different combinations of methods to summarize trees, and to identify potential candidates for pruning.

The main difference of the approach explored here is that it conducts two separate evaluations for unstable taxa, carried out first as the pseudoreplicate datasets are generated and analyzed, and then again on the resulting strict consensus trees. The first analysis aims to detect which taxa are commonly unstable among the most parsimonious topologies derived from a single pseudoreplicate. During each pseudoreplicate this evaluation is independently repeated and the MPTs of each pseudoreplicate are saved to temporary tree files (without collapsing) so that they do not have to be calculated again. The second analysis is conducted by analyzing the taxa that are unstable when comparing the 100 strict consensus trees derived from each of the replicates (fig. 3). This two-step procedure constitutes the major difference with previous approaches and provides a list of potentially unstable taxa that can decrease the support values in resampling measures. The final step consists of determining the optimal subset of unstable taxa that should be pruned to obtain an informative jackknife tree. Once the taxa are identified, they can be pruned from the trees saved for each pseudoreplicate for calculation of the reduced strict consensus, followed by the majority rule consensus. We explore the implementation of this protocol for jackknife analysis, but equivalent implementations for bootstrap and symmetric resampling can be easily run through the use of an argument in the script used here (see below).

**TAXON INSTABILITY WITHIN PSEUDOREPLICATES:** Given that for each pseudoreplicate a

tree search and a strict consensus is performed, the detection of unstable taxa in this context is entirely equivalent to the problem of detecting unstable taxa for producing a reduced consensus. Therefore, we employ for this analysis the IterPCR method (Pol and Escapa, 2009) as implemented in TNT (Goloboff and Szumik, 2015). The taxa detected are in many cases known from highly incomplete specimens and have a large number of missing entries. These taxa, after the deletion of a few characters during the resampling procedures, become unstable in the tree search conducted in the perturbed dataset. Such taxa will create polytomies in the strict consensus tree that will obscure the relationships of the stable taxa in the analysis of a given pseudoreplicate. In our implementation, a taxon is added to a preliminary list of unstable taxa when it is detected as unstable in at least 15% of the pseudoreplicates (although this threshold value can be changed as an argument of the script in TNT).

**TAXON INSTABILITY AMONG PSEUDOREPLICATES:** The second analysis of taxon instability is conducted using the 100 strict consensus trees derived from each of the pseudoreplicates. The taxa that are unstable among these strict consensus trees are in many cases terminals that have a conflicting combination of characters and that are depicted in alternative positions depending on the combination of characters left in the perturbed dataset after the random deletion of certain characters. This procedure is performed using a modified version of the IterPCR script (Pol and Escapa, 2009) in which triplets are regarded as in agreement if they show the same resolution in more than half the topologies analyzed. This differs from the original implementation that counted agreement in triplets only when all the topologies shared the same resolution. This follows the logic of majority rule reduced consensus trees (Wilkinson, 1996) given that this is the method used for obtaining absolute frequencies for the jackknife support analysis. The set of taxa detected as unstable among the strict consensus

trees derived from the 100 replicates is added to a preliminary list of unstable taxa.

**OPTIMAL SET OF UNSTABLE TAXA:** The final step consists of determining which of the taxa in the preliminary list affect the jackknife values in a significant way. Their effect on the resampling support values is then evaluated with the command *prupdn* (implemented by P.A.G.) in TNT ver 1.5 (Goloboff and Catalano, 2016). The preliminary list of unstable taxa is input to this command, which evaluates different combinations of prunes to improve the majority rule consensus. For each combination the command *prupdn* evaluates an optimality function  $E$ :

$$E = ( P + \sum jak ) / ( T - 2 )$$

where *jak* are the jackknife support values of each node of the pruned tree,  $P$  is a penalty for pruning taxa, and  $T$  is the number of taxa with least possible prunings. The penalty factor ( $P$ ) is a function of the number of taxa removed ( $R$ ) and a scaling factor ( $F$ ) that can be adjusted between 0 and 100, with  $P = 100 R (1 - F^2)$ . Using  $F = 1$  results in no penalization for pruning taxa ( $P = 0$ ) and selects the combination that maximizes the sum of support values of the reduced jackknife tree. Using values of  $F > 1$  will penalize the pruning of taxa and values of  $F < 1$  will favor the pruning of taxa. Further options of this command can be found in the TNT documentation (Goloboff and Catalano, 2016).

Note that the *prupdn* command evaluates  $E$  for a large number of combinations of prunes (the number of combinations to try is somewhat reduced by using groups with a high frequency as “separators” [see TNT documentation], but it is still a large number). This is why it is important to produce a preliminary list of taxa potentially affecting support values, found by more approximate methods that do not use an explicit optimality criterion, instead of pruning all combinations of all taxa. In the present implementation, that preliminary list is obtained by applying IterPCR to improve the strict consensus for the results of individual pseudoreplicates, and a

modified IterPCR that evaluates triplet frequency for the majority rule tree of the consensus resulting from all the pseudoreplicates. It would be easy to modify the script used here so that alternative ways to produce preliminary lists of prunes are used (e.g., with the TNT commands *chkmoves*, *prunnelsen*, or *prunmaj*).

**IMPLEMENTATION:** The entire procedure is implemented in *pcrjak.run* (see <http://www.lillo.org.ar/phylogeny/tnt/scripts/pcrjak.run>), a TNT script that takes as input the dataset and performs the jackknife pseudoreplicates, applies the two successive steps of identification of unstable taxa (within and among replicates), and then identifies the optimal set of unstable taxa that can be pruned in order to maximize the function  $E$  (see above).

To run the script, the user needs to call the script giving the data matrix name as the first argument (e.g., *pcrjak datafile.tnt*;). The default option is to use the jackknife resampling method, but this can be changed giving a second argument. For example, *pcrjak datafile.tnt boot*; will perform the same procedure on the data matrix *datafile.tnt* using the bootstrap resampling method.

The outputs of the script are two files. First, a graphical tree (*pcrjak.svg*) that shows the reduced jackknife tree with their absolute frequency values (after ignoring the position of the unstable taxa). Second, a plain text file (*pcrjak.out*) in which the same tree is depicted (in text format) followed by the list of the selected unstable taxa and the lists of taxa that are unstable within and among pseudoreplicates. We recommend using the command-based version of TNT (either for Windows, Mac, or Linux) so that all temporary tree files are deleted automatically.

## RESULTS

Applying the above-described protocol for the TWiG dataset results in the identification of 35 taxa that decrease the jackknife support values. When the positions of these taxa are ignored for constructing the majority rule tree (reduced majority rule consensus tree; Wilkinson, 1996),

a total of 104 nodes are recovered in this tree (in comparison to the 68 nodes present in the jackknife tree when all the taxa are included). The reduced jackknife tree (fig. 4) has an average support value of 77%, showing that numerous clades of coelurosaurian theropods are well supported within the context of the TWiG dataset. The support of the 36 nodes retrieved in the reduced jackknife analysis that are absent from the complete jackknife analyses was obscured only by the alternative positions of unstable taxa.

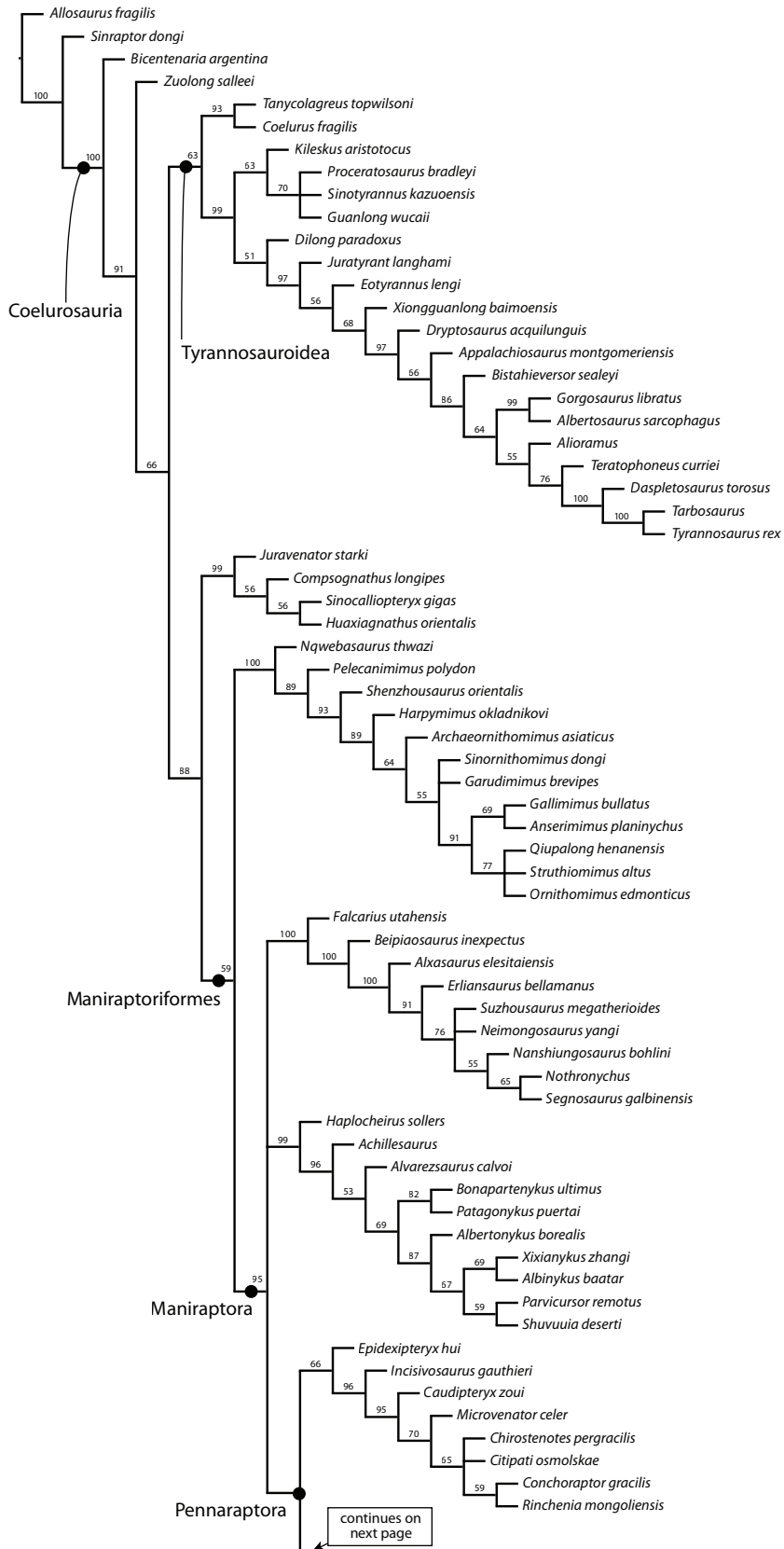
The protocol implemented identified 33 taxa that are unstable within the jackknife pseudoreplicates and 55 taxa that are unstable among the 100 strict consensus of the jackknife pseudoreplicates (table 1). These two sets of unstable taxa are only partially overlapping and share 22 taxa (*Graciliraptor*, *Atrociraptor*, *Utahraptor*, *Saurornitholestes*, *Shanag*, *Austroraptor*, *Hesperonychus*, *Luanchuanraptor*, *Yurgovuchia*, IVPP V22530, *Archaeopteryx* Haarlem, IGM 10044, *Iaceornis*, *Limenavis*, *Liaoningornis*, *Erlikosaurus*, *Enigmosaurus*, *Nanshiungosaurus*, *Ceratonykus*, *Beishanlong*, *Mirischia*, *Velociraptor osmolskae*). The partial overlap among these taxon sets shows that taxa that are unstable within trees found in pseudoreplicates are not necessarily unstable in the strict consensus among pseudoreplicates. Out of the combined set of 66 unstable taxa detected by this procedure, the command *prupdn* found that 35 taxa (table 1) are the ones that maximize the optimality function *E* and therefore more markedly decrease jackknife support values. The composition of these 35 taxa highlights the importance of the two-step procedure for identifying unstable taxa: five of them are unstable only within the trees from each pseudoreplicate, 14 of them are unstable only among the strict consensus derived from each pseudoreplicate, and finally 16 taxa listed above were detected as unstable both within and among pseudoreplicates.

Many of the taxa detected as unstable are known from incomplete specimens, in particular those detected to vary within pseudoreplicates, having an average of 85% of missing

entries. Taxa detected as unstable among pseudoreplicates have an average of 76% of missing entries (slightly larger than the 65% average across the entire dataset). The amount of missing entries, however, does not unequivocally indicate instability, given that several taxa detected as unstable have a percentage of missing entries that is below the overall average for the dataset (15 of the taxa that are unstable among pseudoreplicates and 3 of the taxa that are unstable within). Similarly, there are five taxa with >90% missing entries that are not unstable among jackknife trees. The distribution of missing entries among terminal taxa across time (fig. 5; gray circles) shows that taxa detected as unstable in jackknife trees within (fig. 5; green triangles) or among (fig. 5; blue squares) pseudoreplicates, as well as the subset selected by the *prupdn* command (fig. 5; solid red circles) include some taxa that are fairly well known (e.g., *Ornitholestes*, *Sinovenator*, *Sinosauropteryx*), for which the instability likely derives from character conflict more than from absence of information.

Below we discuss the results that arise from the jackknife analysis based on reduced majority rule consensus for several clades of coelurosaurian theropods, contrasting with those obtained when all the taxa are included in the support trees and when the results of the jackknife analysis are summarized using the majority rule consensus (as in the default option of TNT).

**COELUROSAURIA:** The monophyly of Coelurosauria has high support, but the taxon sampling includes few non-coelurosaurian theropods. *Tugulusaurus* is detected as unstable and causing low supports at the base of Coelurosauria in the jackknife analysis including all taxa (fig. 2). Ignoring this taxon in the reduced jackknife consensus (fig. 4) resolves *Zuolong* as the sister group of the clade formed by Maniraptoriformes and Tyrannosauroida (although still with low support 66%; fig. 4). Multiple subclades within Tyrannosauroida are well supported (>90%) in both the complete jackknife tree and the reduced jackknife trees (figs. 2, 4).



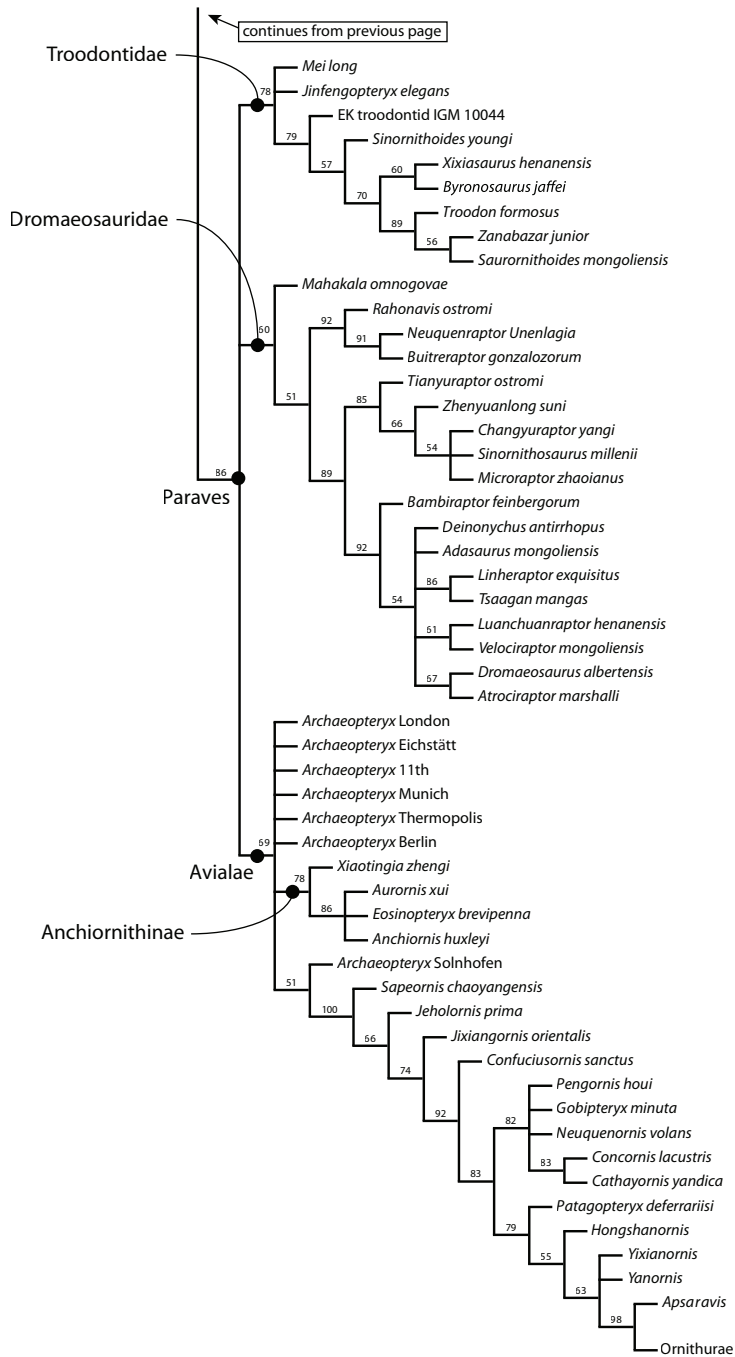


FIG. 4. Absolute frequency jackknife tree for the dataset of Pei et al. (in press) obtained by the script *pcrjak*. *run*, ignoring the alternative positions of 35 taxa detected as unstable in the support analysis that decrease jackknife values.

TABLE 1

**Taxa detected as unstable within and among jackknife pseudoreplicates and those selected by *prupdn* as unstable taxa decreasing support in jackknife support analysis**

Unstable among jackknife pseudoreplications	Unstable within jackknife pseudoreplicates	Selected by <i>prupdn</i> as unstable taxa
–	<i>Acheroraptor</i>	<i>Acheroraptor</i>
<i>Achillobator</i>	–	<i>Achillobator</i>
<i>Adasaurus</i>	–	–
–	<i>Albinykus</i>	–
<i>Archaeopteryx</i> Haarlem	<i>Archaeopteryx</i> Haarlem	<i>Archaeopteryx</i> Haarlem
<i>Atrociraptor</i>	<i>Atrociraptor</i>	–
<i>Austroraptor</i>	<i>Austroraptor</i>	<i>Austroraptor</i>
<i>Avimimus</i>	–	<i>Avimimus</i>
<i>Balaur</i>	–	<i>Balaur</i>
–	<i>Baptornis</i>	–
<i>Beishanlong</i>	<i>Beishanlong</i>	<i>Beishanlong</i>
<i>Buitreraptor</i>	–	–
<i>Byronosaurus</i>	–	–
<i>Ceratomykus</i>	<i>Ceratomykus</i>	<i>Ceratomykus</i>
<i>Changyuraptor</i>	–	–
<i>Compsognathus</i>	–	–
<i>Dakotaraptor</i>	–	<i>Dakotaraptor</i>
<i>Dromaeosaurus</i>	–	–
<i>Enigmosaurus</i>	<i>Enigmosaurus</i>	<i>Enigmosaurus</i>
<i>Erlikosaurus</i>	<i>Erlikosaurus</i>	<i>Erlikosaurus</i>
<i>Graciliraptor</i>	<i>Graciliraptor</i>	<i>Graciliraptor</i>
<i>Hesperonychus</i>	<i>Hesperonychus</i>	<i>Hesperonychus</i>
–	<i>Hesperornis</i>	–
<i>Iaceornis</i>	<i>Iaceornis</i>	–
IGM 10044	IGM 1044	–
<i>Ingenia</i>	–	<i>Ingenia</i>
IVPP V22530	IVPP V22530	IVPP V22530
<i>Jinfengopteryx</i>	–	–
<i>Liaoningornis</i>	<i>Liaoningornis</i>	<i>Liaoningornis</i>
<i>Limenavis</i>	<i>Limenavis</i>	–
<i>Linhenykus</i>	–	<i>Linhenykus</i>
<i>Linheraptor</i>	–	–
–	<i>Lithornis</i>	–
<i>Luanchuanraptor</i>	<i>Luanchuanraptor</i>	–
<i>Mei</i>	–	–
<i>Mirischia</i>	<i>Mirischia</i>	<i>Mirischia</i>
<i>Mononykus</i>	–	<i>Mononykus</i>
<i>Nanshiungosaurus</i>	<i>Nanshiungosaurus</i>	–
<i>Neimongosaurus</i>	–	–
–	<i>Neuquenornis</i>	–
<i>Ornitholestes</i>	–	<i>Ornitholestes</i>

TABLE 1 *continued*

Unstable among jackknife pseudoreplications	Unstable within jackknife pseudoreplicates	Selected by <i>prupdn</i> as unstable taxa
<i>Oviraptor</i>	–	<i>Oviraptor</i>
–	<i>Patagopteryx</i>	–
<i>Rahonavis</i>	–	–
–	<i>Rinchenia</i>	–
<i>Saurornithoides</i>	–	–
<i>Saurornitholestes</i>	<i>Saurornitholestes</i>	<i>Saurornitholestes</i>
<i>Shanag</i>	<i>Shanag</i>	<i>Shanag</i>
<i>Sinornithoides</i>	–	–
<i>Sinosauropteryx</i>	–	<i>Sinosauropteryx</i>
<i>Sinovenator</i>	–	<i>Sinovenator</i>
<i>Sinusoanator</i>	–	<i>Sinusoanator</i>
–	<i>Songlingornis</i>	<i>Songlingornis</i>
<i>Suzhousaurus</i>	–	–
–	<i>Therizinosaurus</i>	<i>Therizinosaurus</i>
<i>Troodon</i>	–	–
<i>Tsaagan</i>	–	–
<i>Tugulusaurus</i>	–	<i>Tugulusaurus</i>
<i>Unenlagia</i>	–	–
<i>Utahraptor</i>	<i>Utahraptor</i>	<i>Utahraptor</i>
–	<i>Vorona</i>	<i>Vorona</i>
<i>Velociraptor osmolskae</i>	<i>Velociraptor osmolskae</i>	<i>Velociraptor osmolskae</i>
<i>Xixiasaurus</i>	–	–
<i>Yi</i>	–	<i>Yi</i>
<i>Yurgovuchia</i>	<i>Yurgovuchia</i>	<i>Yurgovuchia</i>
<i>Zanabazar</i>	–	–

**MANIRAPTORIFORMES:** This clade had low support (62%; fig. 2) and its early-diverging nodes were collapsed in the jackknife analysis including all the taxa, but the reduced jackknife analysis reveals high support (88%) for the node including compsognathids, ornithomimosaurians, and later-diverging coelurosaurians. The reduced jackknife tree also resolves compsognathids as earlier diverging than ornithomimosaurians, although with low support (59%; fig. 4). This improvement is a consequence of the unstable position of *Ornitholestes* and two compsognathids (the highly incomplete *Mirischia* and the better-known *Sinosauropteryx*) that were unstable within many replicates. The monophyly of both Compsognathidae and Ornithomimosauria are highly supported in the reduced jackknife tree (99% and 100%; fig. 4).

The support of Maniraptora is moderate in the jackknife based on the complete taxon set (75%; fig. 2), but increases to 95% in the reduced jackknife tree. This difference is due to the unstable behavior of pennaraptorans of uncertain affinities that are excluded from the reduced jackknife tree (see below). The monophyly of the two earliest-diverging branches (Alvarezsauriidae and Therizinosauria) is also extremely well supported (>90%) in both the complete jackknife tree and the reduced jackknife trees (figs. 2, 4).

**PENRARAPTORA:** The difference in the complete jackknife tree and the reduced jackknife trees in this clade is the largest of the analysis (figs. 2, 4). A high-order polytomy is present when all the taxa are included and poorly supported groups are collapsed. Pennaraptora has moderate support in the reduced tree (78%; fig.

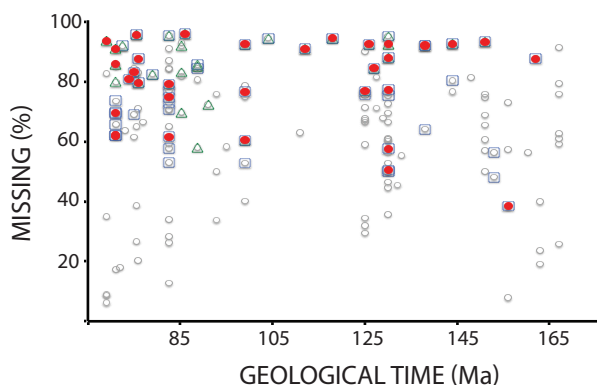


FIG. 5. Distribution of missing data in terminal taxa of the TWiG dataset (Pei et al., in press) through geological time. Each point represents a terminal taxon, vertical axis represents percentage of missing data in the matrix, and horizontal axis represents age of the terminal taxon (in millions of years). The graph represents all terminal taxa (gray circles), taxa identified as unstable within each pseudoreplicate (green triangles), taxa identified as unstable among the strict consensus of pseudoreplicates (blue squares), and subset of taxa identified as unstable by the *prupdn* command (solid red circles).

4) and resolves the scansoriopterygid *Epidexipteryx* and oviraptorosaurians as the sister group of Paraves. Oviraptorosaurians are well supported in the reduced jackknife tree (96%; fig. 4) as well as in the complete jackknife tree (89%; fig. 2). Support within this clade is low even when unstable oviraptorids are excluded (e.g., *Oviraptor*, *Ingenia*), except for the early-diverging position of *Incisivosaurus* within this clade. One of the unstable taxa excluded from the reduced jackknife tree that resolves Pennaraptora is the bizarre scansoriopterygid *Yi qi*, reflecting the conflictive history of hypotheses about scansoriopterygid affinities (Zhang et al., 2008; Turner et al., 2012; Agnolín and Novas, 2013; Godefroit et al., 2013).

**PARAVES:** In the jackknife tree including all taxa, Paraves has very low support, but this clade is retrieved in the reduced jackknife tree with a relatively high nodal support (86%; fig. 4). The relationships between the three main paravian clades (Dromaeosauridae, Troodontidae, Avialae) are weakly supported in the reduced jackknife tree, demonstrating the lack of strong support for the topology retrieved in the MPTs (fig. 1), in which Deinonychosauria is depicted as monophyletic. The lack of support for the rela-

tionships of the three major clades of paravians is not caused by unstable taxa but instead due to character conflict in this dataset.

Dromaeosauridae is collapsed in the jackknife tree including all taxa and has low support (60%; fig. 4) when unstable taxa are pruned from the jackknife trees. The internal relationships of Dromaeosauridae also have relatively low supports, such as the nodes within Microraptoria and within Eudromaeosauria. The only nodes that are well supported in Dromaeosauridae are Microraptoria, Eudromaeosauria, and the node that clusters these two clades (89%–92%; fig. 4). Additionally, the support for Unenlagiinae and for the sister group relationship of *Buitreraptor* and *Unenlagia/Neuquenraptor* are also well supported (>90%; fig. 4) when the unstable *Austroraptor* is excluded. These well-supported dromaeosaurid nodes were completely collapsed in the jackknife tree including all the taxa, so that unstable dromaeosaurids obscured the support of these nodes. These support values are obtained after pruning a total of 11 dromaeosaurids (out of the 31 included in this matrix), identified as unstable by the *prupdn* command (i.e., *Saurornitholestes*, *Graciliraptor*, *Hesperonychus*, *Shanag*, *Zhenyuanlong*, *Dakotaraptor*, *Yurgov-*



*uchia*, *Austroraptor*, *Achillobator*, IVPP V22530, and *Utahraptor*). Dromaeosaurids concentrate the largest number of unstable taxa among all coelurosaurian clades in this dataset, and it is worth noting that some of these taxa are among the most complete of the 31 pruned taxa.

In contrast to Dromaeosauridae, Troodontidae is retrieved as moderately supported (73%; fig. 2) in the jackknife tree when all the taxa are included. The support for this node increases marginally (78%; fig. 4) in the reduced jackknife tree, after the exclusion of *Sinusonasus* and *Sino-venator*. As in the case of Dromaeosauridae, the internal relationships of troodontids have generally low support (56–89%; fig. 4).

Anchiornithinae is recovered in the jackknife tree when all the taxa are included (68%; fig. 2) and when the unstable taxa are pruned (78%; fig. 4). The early-diverging position of *Xiaotingia* is also present in both trees, although pruning the unstable taxa increases the support value from 73% (fig. 2) to 96% (fig. 4). Note that none of the anchiornithines are pruned in the reduced jackknife tree, so this increase is caused by non-anchiornithine taxa that influence the results of the jackknife analysis of this node.

The analysis of Pei et al. (in press) retrieves Anchiornithinae as the sister group of Avialae in the MPTs. The sister group relationship of Anchiornithinae and Avialae occurs with frequency below 50% in the complete jackknife tree, but occurs in the reduced jackknife tree with moderate support (69%; fig. 4). The affinities of anchiornithines are one of the most contentious issues in paravian systematics and vary significantly in previous analyses depending on the taxon and character sampling schemes (see Agnolín and Novas, 2013; Brusatte et al., 2014; Xu et al., 2015, 2017; Foth and Rauhut, 2017; Hu et al., 2018). This conflict is clearly not determined by unstable taxa within this clade but is affected by unstable taxa of other clades (e.g., dromaeosaurids and avialans) and the presence of character conflict (Xu and Pol, 2012).

Finally, Avialae is retrieved only in the reduced jackknife tree with low support (<50%;

fig. 4), which is partially caused by the scoring of *Archaeopteryx* at the specimen level. After pruning the Haarlem specimen that was identified as highly unstable, most specimens of *Archaeopteryx* are in a polytomy at the base of Avialae in the reduced jackknife tree (fig. 4). Previous authors have noted the presence of character conflict related to the affinities of *Archaeopteryx* with early-diverging avialans (Xu et al., 2011; Turner et al., 2012; Xu and Pol, 2012; Godefroit et al., 2013) and this character conflict is present in the analyzed dataset. Combining all the specimens into a single composite terminal taxon does not result in high support for the position of *Archaeopteryx* at the base of Avialae. After combining all the specimens into a single OTU and running the *pcrjak* script the support for Avialae (*Archaeopteryx* + later-diverging birds) is still extremely low (53%). The low support retrieved in the reduced jackknife trees for the position of *Archaeopteryx* at the early-diverging node of Avialae reflects the presence of the character conflict noted by previous authors. The node formed by *Sapeornis* and later-diverging avialans is retrieved as monophyletic with moderately high support (83%; fig. 2) and maximal values are obtained after pruning unstable taxa (100%; fig. 4). Within avialans, some unstable taxa are pruned (e.g., *Vorona*, *Liaoningornis*, *Songlingornis*, *Limenavis*), but these are involved in small polytomies in the complete jackknife tree (fig. 2).

## CONCLUSIONS

The protocol implemented for producing reduced jackknife trees reveals that a large number of coelurosaurian clades are well supported in the latest version of the TWiG dataset (Pei et al., in press) when unstable taxa are pruned from the trees found during the resampling procedure. The unstable taxa that decrease nodal support are detected in both the optimal trees within individual jackknife pseudoreplicates and among the strict consensus of the different jackknife pseudoreplicates. The two-step approach used here for

parsimony analysis is automated in a script for TNT (Goloboff et al., 2008) and differs from previous attempts to identify unstable taxa and to produce well-resolved trees in resampling procedures of nodal support (e.g., bootstrap/jackknife; Wilkinson, 1996; Pattengale et al., 2011; Aberer and Stamatakis, 2011; Aberer et al., 2013).

Previous support analyses on the TWiG dataset resulted in highly collapsed jackknife trees, which were largely influenced by the presence of fragmentary and/or conflictive taxa that hide nodal support for multiple clades. A total of 31 unstable taxa are identified as the most influential for obscuring nodal support. Although some of these taxa are fragmentary, others have a comparatively low amount of missing entries and suggest character conflict is also a cause of their instability. Major clades that are well supported in the reduced jackknife tree include Coelurosauria, Maniraptoriformes, Compsognathidae, Ornithomimosauria, Alvarezsauridae, Therizinosauria, Oviraptorosauria. Other clades, in turn, have moderate to low support even when unstable taxa are pruned from the resampling trees: Maniraptora, Pennaraptora, Paraves, Dromaeosauridae, Troodontidae, Anchiornithinae, and early-diverging nodes of Avialae.

Reduced jackknife analysis as explored here not only reveals well-supported clades but (more importantly) also highlights specific terminal taxa producing low support of surrounding clades (due to character conflict or lack of enough phylogenetic information). This helps to identify taxa or areas of the phylogeny that require further research efforts in data collection based on comparative anatomical studies.

#### ACKNOWLEDGMENTS

This study was possible thanks to the International Pennaraptoran Dinosaur Symposium held at the University of Hong Kong and supported by Kenneth H.C. Fung and First Initiative Foundation. This study was supported by the Proyecto de Unidad Ejecutora from CONICET (PUE0070 to P.A.G). TNT is made freely available thanks to the Willi Hennig Society.

#### REFERENCES

- Aberer, A., and A. Stamatakis. 2011. A simple and accurate method for rogue taxon identification. IEEE BIBM 2011, Atlanta, GA, November 2011.
- Aberer, A., Krompas, D., and Stamatakis, A. 2013. Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. *Systematic Biology* 62: 162–166.
- Agnolín, F.L., and F.E. Novas. 2013. Avian ancestors: A review of the phylogenetic relationships of the theropods Unenlagiidae, Microraptoria, *Anchiornis* and Scansoriopterygidae. Dordrecht: Springer.
- Brusatte, S.L., G.T. Lloyd, S.C. Wang, and M.A. Norell. 2014. Gradual assembly of avian body plan culminated in rapid rates of evolution across the dinosaur-bird transition. *Current Biology* 24 (20): 2386–2392.
- Farris, J.S., V.A. Albert, M. Källersjö, D. Lipscomb, and A.G. Kluge. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12: 99–124.
- Foth, C., and O.W.M. Rauhut. 2017. Re-evaluation of the Haarlem *Archaeopteryx* and the radiation of maniraptoran theropod dinosaurs. *BMC Evolutionary Biology* 17: 236.
- Gauthier, J.A. 1986. Saurischian monophyly and the origin of birds. In K. Padian (editor), *The origin of birds and the evolution of flight*: 1–55. San Francisco: California Academy of Science.
- Godefroit, P., et al. 2013. A Jurassic avialan dinosaur from China resolves the early phylogenetic history of birds. *Nature* 498: 359–362.
- Goloboff, P.A., and S.A. Catalano. 2016. TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics* 32 (3): 221–238.
- Goloboff, P.A., and D. Pol. 2005. Parsimony and Bayesian phylogenetics. In V. Albert (editors), *Parsimony, phylogeny, and genomics*: 148–159. Oxford: Oxford University Press.
- Goloboff, P.A., and M.P. Simmons. 2014. Bias in tree searches and its consequences for measuring group supports. *Systematic Biology* 63 (6): 851–861.
- Goloboff, P.A., and C.A. Szumik. 2015. Identifying unstable taxa: efficient implementation of triplet-based measures of stability, and comparison with Phyutility and RogueNaRok. *Molecular Phylogenetics and Evolution* 88: 93–104.
- Goloboff, P.A., et al. 2003. Improvements to resampling measures of group support. *Cladistics* 19: 324–332.
- Goloboff, P.A., J.S. Farris, and K.C. Nixon. 2008. TNT, a free program for phylogenetic analysis. *Cladistics* 24 (5): 774–786.

- Hu, D.Y., et al. 2018. A bony-crested Jurassic dinosaur with evidence of iridescent plumage highlights complexity in early paravian evolution. *Nature Communications* 9: 217.
- O'Connor, J.M.K., and C. Sullivan. 2014. Reinterpretation of the Early Cretaceous maniraptoran (Dinosauria: Theropoda) *Zhongornis haoae* as a scansoriopterygid-like non-avian, and morphological resemblances between scansoriopterygids and basal oviraptorosaurs. *Vertebrata Palasiatica* 52 (1): 3–30.
- Pattengale N., Aberer, A., Swenson, K., Stamatakis, A., and B. Moret. 2011. Uncovering hidden phylogenetic consensus in large datasets. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)* 8: 902–911.
- Pei, R., et al. In press. Potential for powered flight neared by most close avialan relatives but few crossed its thresholds. *Current Biology*.
- Pol, D., and I.H. Escapa. 2009. Unstable taxa in cladistic analysis: identification and the assessment of relevant characters. *Cladistics* 25: 515–527.
- Simmons, M.P., and J.V. Freudenstein. 2011. Spurious 99% bootstrap and jackknife support for unsupported clades. *Molecular Phylogenetics and Evolution* 61: 177–191.
- Turner, A.H., S.H. Hwang, and M.A. Norell. 2007. A small derived theropod from Öösh, early Cretaceous, Baykhangor Mongolia. *American Museum Novitates* 3557: 1–27.
- Turner, A.H., P.J. Makovicky, and M.A. Norell. 2012. A review of dromaeosaurid systematics and paravian phylogeny. *Bulletin of the American Museum of Natural History* 371: 1–206.
- Wilkinson, M., 1994. Common cladistic information and its consensus representation: reduced Adams and reduced cladistic consensus trees and profiles. *Systematic Biology* 43: 343–368.
- Wilkinson, M. 1996. Majority rule reduced consensus and their use in bootstrapping. *Molecular Biology and Evolution* 13: 437–444.
- Wilkinson, M., and M.J. Benton. 1995. Missing data and rhynchosaur phylogeny. *Historical Biology* 10: 137–150.
- Wilkinson, M., J.L. Thorley, and P. Upchurch. 2000. A chain is no stronger than its weakest link: double decay analysis of phylogenetic hypotheses. *Systematic Biology* 49: 754–776.
- Xu, X., You, H.L., Du, K., and F.L. Han. 2011. An *Archaeopteryx*-like theropod from China and the origin of Avialae. *Nature* 475: 465–470.
- Xu, X., and D. Pol. 2013. *Archaeopteryx*, paravian phylogenetic analyses, and the use of probability-based methods for palaeontological datasets. *Journal of Systematic Palaeontology* 12: 323–334.
- Xu, X., et al. 2015. A bizarre Jurassic maniraptoran theropod with preserved evidence of membranous wings. *Nature* 521: 70.
- Xu, X., et al. 2017. Modular evolution in an asymmetrically feathered troodontid with transitional features. *Nature Communications* 8: 14972.
- Zhang, F.C., Z.H. Zhou, X. Xu, X.L. Wang, and C. Sullivan. 2008. A bizarre Jurassic maniraptoran from China with elongate ribbon-like feathers. *Nature* 455: 1105–1108.

