Likelihood reinstates *Archaeopteryx* as a primitive bird

Michael S. Y. Lee1,2,* and Trevor H. Worthy3

1 South Australian Museum, North Terrace, Adelaide, South Australia 5000, Australia
2 School of Earth and Environmental Sciences, University of Adelaide, South Australia 5005, Australia
3 School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales 2052, Australia

*Author for correspondence (mike.lee@samuseum.sa.gov.au).

The widespread view that *Archaeopteryx* was a primitive (basal) bird has been recently challenged by a comprehensive phylogenetic analysis that placed *Archaeopteryx* with deinonychosaurian theropods. The new phylogeny suggested that typical bird flight (powered by the front limbs only) either evolved at least twice, or was lost/modified in some deinonychosaurs. However, this parsimony-based result was acknowledged to be weakly supported. Maximum-likelihood and related Bayesian methods applied to the same dataset yield a different and more orthodox result: *Archaeopteryx* is restored as a basal bird with bootstrap frequency of 73 per cent and posterior probability of 1. These results are consistent with a single origin of typical (forelimb-powered) bird flight. The *Archaeopteryx*-deinonychosaur clade retrieved by parsimony is supported by more characters (which are on average more homoplasious), whereas the *Archaeopteryx*-bird clade retrieved by likelihood-based methods is supported by fewer characters (but on average less homoplasious). Both positions for *Archaeopteryx* remain plausible, highlighting the hazy boundary between birds and advanced theropods. These results also suggest that likelihood-based methods (in addition to parsimony) can be useful in morphological phylogenetics.

**Keywords:** bird origins; Theropoda; maximum likelihood; Bayesian inference; *Archaeopteryx*; phylogeny

1. INTRODUCTION

The status of *Archaeopteryx* as a primitive (stem, basal) bird has been almost universally accepted since its discovery over 150 years ago, and this iconic creature has occupied centre stage in debates about avian origins and evolution in general [1,2]. A recent study [3] presented the most compelling evidence to date challenging this long-held assumption: addition of an intriguing new dinosaur fossil (*Xiaotingia*) to one of the most comprehensive phylogenetic analyses of theropods removes *Archaeopteryx* from birds and places it with deinonychosaurs. Thus, *Archaeopteryx* becomes just one of a plethora of advanced bird-like dinosaurs [2–6], no more closely related to living birds than (for instance)


Vélociraptor. This phylogeny has important implications for the evolution of many features in early birds, such as the morphology of the skull and the flight apparatus [3,6]. However, it was acknowledged that the new phylogeny required further investigation, owing to weak support (Bremer support of 2 and bootstrap less than 50%; [3]). Also, as with many morphological studies, only parsimony (cladistic) methods were employed.

Here, we show that maximum-likelihood and related Bayesian methods, which are widely used in molecular biology but infrequently in palaeontology and morphology, robustly restore *Archaeopteryx* as a basal bird. Such likelihood-based-based methods are widely [7], but not universally [8], thought to be superior to parsimony if, for instance, some taxa are on disproportionately long branches, or particular characters have evolved much more rapidly than others. Such complexities are inherent in palaeontological data, with stratigraphically young and/or rapidly evolving taxa at the end of long branches, and certain characters evolving much more rapidly than others [9]. In particular, where one topology is supported by many fast-evolving (homoplasious) characters and another by fewer but more conservative characters, equal-weighted parsimony might favour the first solution, whereas likelihood methods might favour the second solution.

The divergent retrieved positions of *Archaeopteryx* (with birds using likelihood, with deinonychosaurs using parsimony) indicate that a dramatic reinterpretation of bird origins is not yet required, and demonstrate the potential utility of likelihood-based phylogenetic methods in morphology.

2. METHODS AND TERMINOLOGY

Avialae *sensu* [3] refers to the most-inclusive clade-containing modern birds (exemplified by *Passer domesticus*) but not deinonychosaurs (exemplified by *Dromaeosaurus albertensis*). Aves is often applied to a more restricted node-based clade, e.g. that containing *Archaeopteryx* and modern birds (see [2,4,5]). In the current phylogenetic context (figure 1), the above meanings of Avialae and Aves contain the same set of (known) taxa; we use the vernacular term ‘birds’ to refer to these taxa.

The data matrix of Xu et al. [3] was analysed using parsimony, maximum likelihood (RaxML) and Bayesian inference; formatted datasets with commands are appended in electronic supplementary materials, S1–S3. To confirm the published tree, the data were analysed in PAUP [10], a different parsimony program to that originally used. The following commands were used to sample multiple tree islands but avoid memory overflow on any large single island: hsearch addseq = random nreps = 100 nchuck = 1000 chucksize = 1. Strict and majority-rule consensus trees were computed from the most-parsimonious trees.

The data were also analysed with maximum likelihood and Bayesian inference. Both analyses used the Lewis [11] stochastic model with a gamma parameter to account for rate variation across traits; *Allosaurus* was set as the outgroup, and multi-state characters were unordered (as per the original analysis [3]). Likelihood analyses (figure 1) employed RaxML v. 7.2.8 [12], with 1000 bootstraps followed by a maximum-likelihood search: commands employed were -f a -m MULTIGAMMA-K MK -#1000. As RaxML cannot accept
polymorphic cells for morphological data, these were changed to '?' for the RaxML runs. This affected only eight out of 33 286 cells (0.024%) and was expected to have almost no influence on the resultant topology; parsimony analyses were performed with the polymorphic cells altered to '?' to test this assumption.

Bayesian analyses (figure 2) employed MrBayes v. 3.1 [13], with coding set to variable (to reflect lack of invariant characters in matrix). Four independent runs were used, each consisting of four (one cold and three heated) chains of 10 million generations, sampling every 2000 generations, with a majority-rule consensus. The first 40 per cent of each chain was discarded as burn-in. Convergence across runs in topology and parameters, respectively, was confirmed by (i) split frequencies which had a low standard deviation of approximately 0.01 [13] and were highly correlated [14] (figure 2b), and (ii) effectively identical distributions of parameters which all had high effective sample sizes exceeding 1000 [15].

To evaluate whether differential treatment of fast-evolving (homoplasious) characters might be driving the different topologies retrieved by parsimony, likelihood and Bayesian analyses, pairwise comparisons of trees from these three methods were performed using PAUP. The single best likelihood tree was used, along with the resolved majority-rule consensus trees from parsimony and Bayesian analyses (as characters should not be optimized on unresolved strict consensus trees). The characters which differentially supported each tree and their average consistency and retention indices [16] were ascertained. Analogous calculations are not possible in current versions of RaxML and MrBayes.

3. RESULTS AND DISCUSSION

Analysis using PAUP confirmed the heterodox position of Archaeopteryx with deinonychosaurans retrieved by parsimony [3], exactly replicating the strict consensus tree, and optimal tree length of 1404, found in the original study using a different program. While many more most-parsimonious trees were found (99 000), the original study [3] did not claim to exhaustively sample MPTs, and the true number is likely to be much higher than even the number found here.

Figure 1. Archaeopteryx robustly reinstated as the most basal bird by maximum likelihood (RaxML tree, log L = −7608.41). Birds (including Archaeopteryx) is supported by likelihood bootstrap of 73%.
The other three runs, using AWTY [14]; all splits occurred at similar frequencies across all runs.

Harmonic mean log $L$ majority-rule consensus shown; remainder of tree was also very similar to the likelihood tree in figure 1. Harmonic mean this more orthodox grouping of birds, rather than with deinonychosaurs. Notably, Opteryx and the very similar Wellnhoferia grouped with one striking difference (figures 1 and 2) produced a similar overall topology to parsimony, but unlikely to have biased those results.

Changing the eight polymorphic cells to ‘?’ (question mark) resulted in exactly the same consensus topology (i.e. all eight polymorphic cells were analytically equivalent to ‘?’). Thus, the necessary recoding implemented in the RaxML analysis is unlikely to have biased those results.

The maximum-likelihood and Bayesian analyses produced a similar overall topology to parsimony, but with one striking difference (figures 1 and 2a): Archaeopteryx and the very similar Wellnhoferia grouped with birds, rather than with deinonychosaurs. Notably, this more orthodox grouping of Archaeopteryx and birds has at least moderate support (73% ML bootstrap, 1 Bayesian posterior probability). In contrast, the heterodox grouping of Archaeopteryx with deinonychosaurs in the parsimony analysis was acknowledged to be weakly supported [3]. The positions of Xiaotingia and Anchiornis are poorly resolved; they group weakly with troodontids among deinonychosaurs [17] (figure 1) and if so, are no longer part of Archaeopterygidae sensu Xu et al. [3].

Table 1. The character evidence favouring the parsimony, maximum likelihood and Bayesian trees. Pairwise comparisons using parsimony showing the number of characters ($n$) which differentially favour one tree over another, and the mean consistency (ci) and retention (ri) indices for those suites of characters (averaged over both trees being compared).

<table>
<thead>
<tr>
<th>comparison</th>
<th>n chars</th>
<th>average ci</th>
<th>average ri</th>
<th>n chars</th>
<th>average ci</th>
<th>average ri</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) MP versus (2) ML</td>
<td>64</td>
<td>0.312</td>
<td>0.666</td>
<td>49</td>
<td>0.354</td>
<td>0.728</td>
</tr>
<tr>
<td>(1) ML versus (2) Bayesian</td>
<td>69</td>
<td>0.31</td>
<td>0.677</td>
<td>51</td>
<td>0.369</td>
<td>0.739</td>
</tr>
<tr>
<td>(1) ML versus (2) Bayesian</td>
<td>12</td>
<td>0.253</td>
<td>0.586</td>
<td>5</td>
<td>0.493</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Changing the eight polymorphic cells to ‘?’ (question mark) resulted in exactly the same consensus topology (i.e. all eight polymorphic cells were analytically equivalent to ‘?’). Thus, the necessary recoding implemented in the RaxML analysis is unlikely to have biased those results.

The maximum-likelihood and Bayesian analyses produced a similar overall topology to parsimony, but with one striking difference (figures 1 and 2a): Archaeopteryx and the very similar Wellnhoferia grouped with birds, rather than with deinonychosaurs. Notably, this more orthodox grouping of Archaeopteryx and birds has at least moderate support (73% ML bootstrap, 1 Bayesian posterior probability). In contrast, the heterodox grouping of Archaeopteryx with deinonychosaurs in the parsimony analysis was acknowledged to be weakly supported [3]. The positions of Xiaotingia and Anchiornis are poorly resolved; they group weakly with troodontids among deinonychosaurs [17] (figure 1) and if so, are no longer part of Archaeopterygidae sensu Xu et al. [3].

Examination of the character support for the parsimony, likelihood and Bayesian trees suggests a possible explanation for the divergent placement of Archaeopteryx. ‘Unweighted’ parsimony remains the most common method of analysing morphological data, and treats changes across all branches and across all characters equally [18]. In contrast, the likelihood and Bayesian models used here preferentially concentrate homoplasy on longer branches and/or faster characters (characters are assigned to rate categories following a discretized gamma distribution, using information intrinsic to the dataset). Examination of the characters which differentially support one tree over another is illuminating (table 1). More characters support the parsimony tree over the likelihood tree than vice versa (64 versus 49). However, the characters supporting the parsimony tree are generally more labile than those supporting the likelihood tree (ci 0.312 versus 0.354, ri 0.665 versus 0.728). The same pattern is found when the parsimony and Bayesian trees are compared (69 versus 51; ci 0.310 versus 0.369, ri 0.677 versus 0.729). This is consistent with the difference in tree topologies being at least partially driven by parsimony preferring trees supported by more (but on average more homoplasious) characters, and likelihood and Bayesian analyses preferring trees supported by fewer (but on average more conservative) characters.
As expected, very few characters differentially support the (very similar) likelihood and Bayesian trees (12 versus 5), so not much weight should be put on the differing average ci and ri of these small suites of characters.

The likelihood and Bayesian topology is supported by several derived characters which are unique and unversed (but which are homoplasious on the parsimony tree). These include traits unique to Archaeopteryx and other birds, such as even spaced anterior maxillary teeth (character #89) and other birds, such as anterior caudal vertebrae with distally tapering transverse processes (#373). Conversely, some of the similarities between Archaeopteryx and deinonychosaurs, interpreted in the parsimony analysis as shared derived characters [3], need to be reinterpreted as primitive characters of all birds and deinonychosaurs (paravians), secondarily lost in birds above Archaeopteryx and Weltnhoferia. These include hyperextendibility of the second (= first functional) toe (character #323), a posterior process on the ischium (#334) and a distally located obturator process (#167).

The reinstatement of Archaeopteryx as a basal bird in both likelihood-based analyses contradicts macroevolutionary inferences that relied (at least partly) on the shift of Archaeopteryx into deinonychosaurus. For instance, it does not support the proposal that robust skulls and herbivory are primitive for birds [3]. It also has important implications for the evolution of flight. Powered, forewing-only flight typical of modern birds has been generally inferred to be present only in Archaeopteryx and other birds [2,5,20]. The flight capabilities of some deinonychosaurs remain contentious: the most proficient, Microraptor gui, used both the forelimbs and hindlimbs, and has been argued to be a glider [21] and asymmetric pennaceous feathers indicating aerodynamic ability are known only in dromaeosaurs outside of birds [22]. The likelihood-based phylogenies are consistent with a single origin of typical forewing-driven flight with no losses in basal birds (i.e. the taxa here sampled); in contrast, the parsimony tree implies either minimally two origins (in Archaeopteryx, Weltnhoferia, and in ‘true’ birds), or an earlier origin at the base of Paraves (with subsequent reduced flying ability or modifications to four-winged locomotion in various deinonychosaurus; [17,21]).

The results here do not demonstrate unequivocally that Archaeopteryx belongs with birds rather than with deinonychosaurus; bootstrap support is only approximately 73 percent, while the posterior probability of 1 is tempered by arguments that Bayesian inference can greatly overestimate support if the models implemented are inadequate [8,23]. Resolution of the precise position of Archaeopteryx will likely require more empirical data such as new fossils or novel characters. The alternative placements of Archaeopteryx highlight the closure of the morphological gap between birds and theropod dinosaurs [2–6,17,22], and also suggest that dramatic reinterpretations of early bird evolution are not yet required. Furthermore, the more generally accepted tree, found with likelihood-based methods (and the strong contrast with parsimony), suggests that likelihood-based phylogenetic methods should be used more often in palaeontology and morphology [24], especially for large datasets [3] that have sufficiently large numbers of characters to allow critical parameters in likelihood models to be robustly estimated.

We thank the Australian Research Council linkage grant scheme for funding, Xing Xu and two anonymous referees for comments, and Sam Morrison and the Australian Research Collaboration Service for use of cloud computing.


