Quantifying historical trends in the completeness of the fossil record and the contributing factors: an example using Aves

PersonIdentityServiceImpl

Quantifying historical trends in the completeness of the fossil record and the contributing factors: an example using Aves

Daniel T. Ksepka and Clint A. Boyd

Abstract.—Improvements in the perceived completeness of the fossil record may be driven both by new discoveries and by reinterpretation of known fossils, but disentangling the relative effects of these processes can be difficult. Here, we propose a new methodology for evaluating historical trends in the perceived completeness of the fossil record, demonstrate its implementation using the freely available software ASCC (version 4.0.0), and present an example using crown-group birds (Aves). Dates of discovery and recognition for the oldest fossil representatives of 75 major lineages of birds were collected for the historical period ranging from 1910 to 2010. Using a comprehensive phylogeny, we calculated minimum implied stratigraphic gaps (MIG range) across these 75 lineages. Our results show that a reduction in global MIG values of 1.35 Ga (billion years) occurred over the past century in avian paleontology. A pronounced increase in the average rate of global MIG reduction is noted in the post-1970s interval (290.5 Myr per decade) compared to the pre-1970s interval (31.9 Myr per decade). Although the majority of the improvement in the fossil record of birds has come from new discoveries, substantial improvement (~22.5%) has resulted from restudy and phylogenetic revision of previously described fossils over the last 40 years. With a minimum estimate indicating that at least 1.34 Gyr of gaps remain to be filled between the predicted and observed first appearances of major lineages of crown Aves, there is much progress to be made. However, a notable tapering off in the rate of global MIG reduction occurs between 1990 and 2010, suggesting we may be approaching an asymptote of oldest record discoveries for birds. Only future observations can determine whether this is a real pattern or a historical anomaly. Either way, barring the discovery of fossils that substantially push back the minimum age for the origin of crown-clade Aves, new discoveries cannot continue to reduce global MIG values at the average post-1970s rate over the long term.

Daniel T. Ksepka. Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Campus Box 8208, Raleigh, North Carolina 27695, and Department of Paleontology, North Carolina Museum of Natural Sciences, Raleigh, North Carolina 27695, U.S.A. E-mail: ksepka@gmail.com
Clint A. Boyd. Jackson School of Geosciences, University of Texas at Austin, 1 University Station C1100, Austin, Texas 78712, U.S.A. E-mail: clintboyd@stratfit.org

Introduction

Estimating the completeness of the fossil record has long been a central area of research in paleontology (Newell 1959; Valentine 1970; Raup 1972; Maxwell and Benton 1990; Norell and Novacek 1992; Benton 1999; Foote and Sepkoski 1999; Peters and Foote 2001). Quantitative studies suggest that the quality of the fossil record is improving (Benton and Storrs 1994; Benton et al. 2000; Fara and Benton 2000; Wills et al. 2008), and one avenue of inquiry involves seeking the patterns and sources of these advances. The true extent of the fossil record will never be fully known, because we can never be confident in having discovered and correctly identified all recoverable fossils. In reality, paleontologists frequently study trends in the perceived quality of the fossil record, comparing observed data against some baseline estimate of the complete record. One major approach to estimating completeness focuses on inferring the percentage of true species diversity that has been discovered through means such as collection curves (e.g., Maxwell and Benton 1994; Bleiweiss 1998; Kalmar and Currie 2010). A second major approach involves estimating the length of stratigraphic gaps in the record of a group. This may be accomplished by applying metrics that quantify the congruence between the stratigraphic record and the topology of a proposed phylogeny (Gauthier et al. 1988; Norell and Novacek 1992; Huelsenbeck 1994; Siddall 1998; Wills 1999; Angielczyk and Kurkin 2003) or calculating the stratigraphic

gaps implied by divergence dates derived from molecular sequence-based analyses (Clarke et al. 2007; Marjanovic and Laurin 2007).

Improvement in the perceived completeness of the fossil record of a given clade through time can be attributed to two primary phenomena: the discovery of new fossils and the recognition of previously overlooked records (typically through phylogenetic revision of known fossils). However, the differential impact of these phenomena on our understanding of the fossil record has never been quantitatively explored. Here, we evaluate the relative completeness of the fossil record of the clade Aves using a new methodology, explore trends over the last 100 years, and compare the differential impact the discovery of new fossils had in driving these patterns versus phylogenetic revision of known material. In this study, we ask several questions: What is the pattern of gains in completeness of the fossil record of birds over the past century? What percentage of these gains has come from new discoveries and what percentage has come from restudy of previously described fossils? Do the data at hand suggest that significant advances will continue to be made on one or both fronts?

Background.—A common misconception that permeated twentieth-century reviews of the fossil record was portrayal of the avian record as meager, characterized by long gaps and comprising mainly small scraps of delicate, hollow bones. Perpetuation of this myth was fueled by both the relatively small historical number of avian paleontologists (versus, for example, fossil mammal workers [Olson 1985]) and the later adoption of cladistic methodologies by ornithologists (see Cracraft 1980). These factors contributed to a status quo in which many fossils were misidentified, resulting in apparent large gaps in the records for many avian clades despite the fact that unrecognized older fossils belonging to those clades had indeed already been recovered. Poor attempts at resolving the affinities of formally described avian fossils over the first half of the twentieth century also contributed to this stasis by leaving many fossils in monotypic groups of unspecified higher relationships or ''form families'' with

no concrete taxonomic standing. This state of affairs caused Stresemann (1959: p. 270), in his famously pessimistic appraisal of the state of avian phylogeny, to lament that ''as far as birds are concerned, there is virtually no paleontological documentation which has revealed such information on the phylogeny as has been the case with the other classes of vertebrates.'' Even more fossils lay languishing in collections unstudied for decades after their initial discovery. Olson (1985: p. 80) summarized this situation when he remarked that ''the paleontological record of birds is already extensive and highly informative; the eventual exposition of the thousands of undescribed specimens already in museums and under study will make it even more so in the near future.''

In recent years, avian paleontology has undergone a ''renaissance'' of sorts (Mayr 2007), putting to rest the portrayal of the avian fossil record as meager and lacking. This victory is due in equal measures to new discoveries and to carefully considered reviews by the past two generations of paleornithologists. Few would now dismiss the fossil record of birds as poor and most would agree that we are in the midst of a period of accelerated discovery. Until now, however, quantitative measures of the completeness of the fossil record of birds have been restricted primarily to evaluating trends in the number of named species.

Minimum Implied Gaps.—Summing the total length of inferred ghost lineages (Norell 1992), or minimum implied gaps (MIGs; Benton 1994; Wills 1999), is one method for measuring completeness in the fossil record. Logically, if a taxon is present at a given time, its sister taxon must also have evolved by that time (Norell 1992). If the appearances of two sister taxa in the fossil record are not simultaneous, then the inferred ghost lineage, or MIG, is equal to the difference between each taxon's oldest known record (OKR; Walsh 1998). An inherent assumption made when measuring MIG is the absence of ancestors in the fossil record. This assumption is both implicit in our own method and also a practical necessity, because ancestors cannot be identified by cladistic analyses.

Standard comparison of MIG values is accomplished by holding the age data and included set of terminal taxa constant while varying the tree topology between analyses (Pol et al. 2004). For the purposes of this investigation, we instead held the included set of terminal taxa and the tree topology constant while varying the age data between analyses to reflect the ages of the OKRs during different historical intervals. We report our results in raw MIG values rather than stratigraphic consistency metric scores (e.g., MSM* or GER) because the latter normalize MIG values relative to the maximum and/or minimum possible MIG value, both of which are affected by varying the age data between analyses.

Sampled Fossil Record versus Recognized Fossil Record.—As mentioned above, improvement in the perceived completeness of the fossil record may be achieved through two different routes that both affect the resulting MIG values. The first of these is discovery of new fossils. For the purposes of this study, ''discovery'' does not refer to the date a fossil is recovered in the field, which is often unreported, but instead to the date it enters the scientific literature. Obviously, new discoveries can fill in gaps in the fossil record of a group, and thus reduce the MIG for that lineage. Because lengthy gaps imply that substantial morphological evolution has taken place, fossils partially filling such gaps are potentially transformative to our understanding of key evolutionary transitions. Such finds are thus often greeted with much fanfare, such as the comparatively late first discoveries of fossil hagfish (Bardack 1991) and lampreys (Bardack and Zangerl 1968). However, new discoveries are not the only method of increasing perceived completeness of the fossil record. A second event that may reduce global MIG values is the systematic revision of previously described fossils, which may identify the phylogenetic affinities of a fossil for the first time or result in the transfer of a fossil from one lineage to another. Such advances may represent the fruit of increased anatomical scrutiny, further preparation of specimens through new techniques, or revelation of the affinities of previously problematic fossils through phylogenetic analysis.

When exploring the differential impact of new discoveries versus taxonomic revision, it is useful to differentiate between the sampled and recognized fossil record of a lineage. In essence, the ''Recognized Fossil Record'' considers the often imperfect data paleontologists had during a given historical time interval, whereas the ''Sampled Fossil Record'' considers the data at hand for the same interval in light of modern understanding. Because we are looking at historical trends, the present state of knowledge can be taken as a baseline. For the purposes of this study, the Sampled Fossil Record for a given historical interval comprises all fossils that had been discovered by that interval, regardless of whether they were correctly systematically identified during that interval. In contrast, the Recognized Fossil Record for an interval includes only correctly identified fossils. Thus, the Recognized Fossil Record recreates the state of knowledge for a past historical interval. In a hypothetical situation where every fossil was, when first described, correctly referred to the lineage it is currently assigned to (i.e., no errors in systematic placement ever occurred), the Sampled Fossil Record and Recognized Fossil Record would always be equivalent. However, this is often not the case. In instances where fossils are described at one point in time, but not correctly referred to a lineage until years later, the Sampled Fossil Record and Recognized Fossil Record will differ. Specifically in the context of our study, the value of global MIG calculated using the Recognized Fossil Record may remain higher than the value calculated from the Sampled Fossil Record over the historic interval between the original description and the subsequent systematic revision of a fossil.

Examples of the phenomenon of the Recognized Fossil Record lagging behind the Sampled Fossil Record are common in avian paleontology. For example, von Meyer (1839, 1844) originally described the Oligocene fossil bird *Protornis glarniensis* as a possible songbird (Passeriformes). More than a century later, Olson (1976) demonstrated that the species is at least a stem member of Momotidae, a group

that up to that point had no Recognized Fossil Record. Thus the Sampled Fossil Record of total group Momotidae in 1844 would extend back to ca. 30 Ma, but the Recognized Fossil Record of Momotidae would have been nonexistent in that same year. Only in 1976 would the Recognized Fossil Record and Sampled Fossil Record for this lineage achieve equivalence. Exploring the historical pattern of global MIG values calculated from the Sampled Fossil Record allows the effects of new fossil discoveries on our understanding of the fossil record to be quantified and plotted over time. Moreover, calculating the differences in global MIG values obtained from the Recognized and Sampled Fossil Records allows the cumulative effects of taxonomic referrals (both positive and negative) to be determined for any given historical interval.

Methods

Construction of Temporal Data Sets.—We constructed two data sets for this investigation. The first data set records the historical pattern of description and subsequent taxonomic revision of the oldest recognized fossil for each of the 75 sampled clades (Recognized Fossil Record data set) and the second data set records when the currently recognized oldest fossil for each clade was originally described, regardless of its original or subsequent systematic referral (Sampled Fossil Record data set). In order to create each data set, we surveyed the paleontological literature for the past century, collecting the OKRs for 75 major clades that correspond to all extant ordinal-level clades and most major subclades of Aves. Fossil discovery and recognition dates were recorded for the contemporary OKR within each of ten decade-long intervals spanning the last century (1910–2010). As mentioned above, although a multi-year gap often exists between the discovery and description of a fossil, the date of discovery is not always reported; therefore, we adopt the date of the first published account of a fossil as the date of discovery. Each terminal taxon was assigned to one of 27 pre-defined age bins that each span 2.5 Myr, with the oldest temporal bin ranging from 65.0 to 67.5 Ma. A 28th age bin was erected for taxa with no fossil record in

order to distinguish them from taxa with a fossil occurrence in the 10 Ka to 2.5 Ma interval. In cases where the uncertainty surrounding the age of a fossil spanned multiple temporal bins, it was assigned to the midpoint of the possible range. Although failing to consider each OKRs full range of temporal uncertainty will slightly affect the resultant MIG values for each historical bin (Pol and Norell 2006), given the large spans of time being worked with in this study (i.e., global MIG values measured in billions of years), the overall effect would not override the broader patterns discussed below. The resulting data sets are presented in the Appendix in the online supplemental material.

Phylogenetic Framework.—We used the phylogeny of Hackett et al. (2008) to generate time-calibrated cladograms (chronograms) for calculating MIG values (see below). The results of the Hackett et al. (2008) study represent one of the only phylogenetic hypotheses for all major extant avian clades based on simultaneous analysis of all included taxa. Importantly, this study also provides a fully resolved tree, which reduces the number of factors contributing to the resulting error range and makes computation of MIG values more tractable. We note that although choosing an alternate phylogeny of Aves would alter global MIG calculated by our method, it would not affect the difference in MIG calculated between two given historical bins. Furthermore, many of the longest reconstructed ghost lineages occur along branches stemming from uncontroversial nodes and so would retain identical lengths under alternate phylogenetic hypotheses.

The cladogram used in our study collapses the individual species exemplars from the Hackett et al. (2008) phylogeny into 75 higher level taxa. Lineages corresponding to all extant avian orders were included. In cases where the monophyly of a traditional order is doubtful (e.g., ''Caprimulgiformes,'' ''Coraciiformes,'' ''Gruiformes''), we included lineages corresponding to the constituent monophyletic clades formerly included in the sundered order. For example, we included separate terminals for five clades (Steatornithidae, Podargidae, Nyctibiidae, Caprimulgidae, and

Aegothelidae) formerly united in ''Caprimul-giformes'' but now recognized to form a paraphyletic assemblage with respect to Apodiformes (Mayr 2002, 2009a; Cracraft et al. 2004; Ericson et al. 2006; Hackett et al. 2008). Taxonomic sampling is limited partially by the inclusiveness of the resolved phylogenies available for Aves (i.e., coverage across Aves is at the order to family level for most clades), and partially by our ability to place fossils phylogenetically. Furthermore, most relevant fossils have been placed as stem representatives of extant ordinal- or family-level clades, which precludes examining these data at a lower taxonomic level.

A handful of extinct clades were excluded in tabulating the OKRs of the extant clades considered here. We considered Pelagornithidae, Dromornithidae, Gastornithidae and Plotopteridae to be of uncertain higher affinities for the purposes of our analysis, given the ongoing controversy over their phylogenetic relationships (Olson and Hasegawa 1979, 1996; Olson 1985; Andors 1992; Murray and Vickers-Rich 2004; Bourdon 2005; Mayr 2005, 2009b). However, we note that because these groups have been allied primarily with lineages with early OKRs (Galloanserae, Sphenisciformes, Suloidea), their inclusion would have little or no effect on our calculations under most proposed placements.

In tabulating Sampled Fossil Record and Recognized Fossil Record in this study, we assumed that the most recent phylogenetic placements of the fossils we considered are correct. Thus, Sampled Fossil Record and Recognized Fossil Record will always be equivalent in the last historical bin (2010 in our study). Because the effects of future phylogenetic revisions of these fossils cannot be known in advance, taking the present-day recognized data set as a baseline is the most reasonable way to facilitate historical comparisons.

Calibrating Observations of MIG across Historical Time.—Frame of reference is important when analyzing historical trends in the fossil record. Paleontologists often measure completeness from time-calibrated cladograms, which minimize implied gaps (e.g., Kerr and Kim 2001; Angielczyk and Kurkin
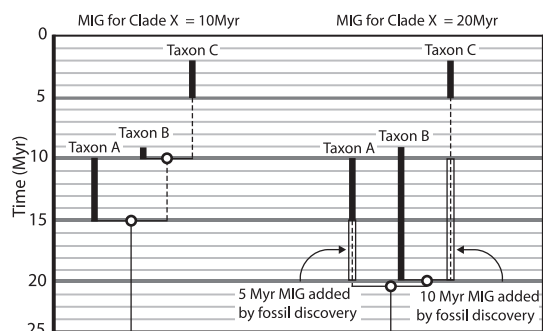

FIGURE 1.   Example illustrating how new fossil discoveries can increase total minimum implied gap (MIG) for a clade. At left, a cladogram depicts the relationships and stratigraphic ranges (black bars) of three taxa in clade X. Implied gaps are depicted with dashed lines. At right, the effects of a new fossil discovery extending the range of taxon B are shown. The new fossil reduces the MIG for the clade containing taxa B and C. However, the new fossil also implies additional, previously unrecognized gaps in the fossil records of taxa A and C, resulting in an overall increase in MIG for clade X.

2003; Jeffery and Emlet 2003). These time-calibrated cladograms combine tree topology and the ages of taxa with more complete records to estimate the MIG of taxa with less complete records. Sister taxa by definition diverged simultaneously from their common ancestor. Thus, unless the oldest records of both sister lineages are equivalent in age, any extension of the fossil record of the older sister taxon will add to the MIG of the younger sister taxon. This presents an issue for analyzing historical trends, because it is easy to contrive a situation where new discoveries increase the global MIG for a cladogram. For example, the discovery of the new fossil in Figure 1 would extend the range of taxon B by 10 Myr, representing an increase in the overall completeness of the true fossil record. However, this same discovery would also imply previously unrecognized additional gaps of 10 Myr in the fossil record of taxon A and of 5 Myr in the fossil record of taxon C. Thus, the new fossil discovery would actually result in net *increase* of 10 Myr in the MIG for the groups under consideration.

In order to provide a baseline for historical comparisons and correct for artificial increases in MIG across time, we used the 2010 Sampled Fossil Record data as a proxy of the ''true'' fossil record when calculating MIG. This was accomplished by calculating MIG values from
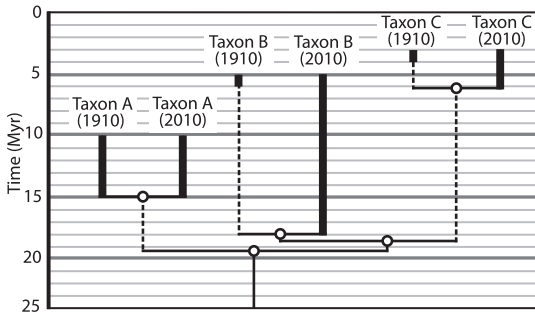
FIGURE 2. Example illustrating the method used by this study for preventing new fossil discoveries from artificially increasing global minimum implied gap (MIG) values. Each terminal branch is duplicated (i.e., a sister taxon is added), with one of the resultant terminals assigned the currently recognized oldest known record (OKR) for that taxon, and the other assigned the OKR as it was known during the historical bin being investigated (e.g., 1910). This sets the present-day fossil record as the baseline from which global MIG values for all historical bins are calculated. Known stratigraphic ranges for terminal taxa are depicted by thickened black bars and implied gaps are shown by dashed lines.

a tree in which each terminal branch is duplicated (i.e., a sister taxon is added to each terminal taxon; see Fig. 2). One of the duplicated branches was assigned the 2010 OKR and the other was assigned the OKR for the historical bin being considered (e.g., 1910). This strategy ensures that discoveries of older fossils always reduce MIG.

It should be noted that this practice makes the assumption that the currently known fossil record includes the oldest discoverable fossils of each group under consideration. This is highly unlikely, because older records of major avian clades continue to be published. However, this method provides a necessary baseline for calculating the amount of global MIG reduction accomplished over the last 100 years. Fortunately, this method of standardizing does not affect the difference in MIG calculated between two given historical bins. If older fossils are discovered subsequent to our analyses, the *amount* of MIG calculated over the avian tree for each historical bin will increase, but the *difference* between the MIG values for two given historical bins will not. Correspondingly, the slope of the curves we present below would remain the same though the y-intercept would change. Thus, the working assumption that the known fossil record includes the oldest

discoverable fossils keeps the frame of reference standardized without modifying the results in terms of trends.

Finally, we note that our data set exclusively considers gaps that occur between the phylogenetically predicted first appearance of a clade and that clade's OKR. However, gaps in the fossil record may occur throughout a clade's stratigraphic range. New fossil discoveries can reduce MIG for a lineage even though large internal gaps remain. Thus, a group with a very short MIG may still have a poor overall fossil record. Because the vast majority of fossil bird species are known from a single fossil horizon, data for birds are more amenable to methods that utilize MIG than those that attempt to consider gaps throughout a clade's entire stratigraphic range.

Software Implementation.—Calculation of MIG range values was conducted using the software program ASCC version 4.0.0 (Boyd et al. 2011), which is designed to simplify the calculation of stratigraphic consistency metric values. The tree topology was imported into the program and displayed in the Tree Data window. The base of the tree was highlighted and the option ''add sister taxon to subtree'' was selected. This has the effect of adding a sister taxon to every terminal taxon in the analysis (see above). The Taxon Data window was then used to define the age of the oldest known record for each terminal taxon. Each analysis was run for 1,000,000 replications by selecting the Calculate Scores option under the Tools menu. In each replicate, an age is randomly selected for each terminal taxon from the designated age bin and a value of MIG is calculated and saved. Once the analysis is complete, the program automatically combines the highest and lowest observed values into the final range score, which describes the full range of recovered values. This process was repeated for each of the 11 time bins (1910–2010) in both the Sampled Fossil Record and Recognized Fossil Record data sets. The results of all 22 analyses can be found in Table 1.

Results

General Trends in the Recognized Fossil Record Data Set.—Global MIG values for 1910 average 2.73 Gyr, declining over the next century to the

TABLE 1.   Resulting minimum and maximum values of MIG for each historical bin for both the 'Recognized Fossil Record' and 'Sampled Fossil Record' datasets. The 'new discoveries' column indicates how much MIG reduction should have occurred during each historical bin due solely to new discoveries. The 'cumulative effect of taxonomic work' column describes how taxonomic work on new and previously known fossils either underperformed (negative values) or outperformed (positive values) the expected reduction in global MIG values due to new discoveries. All MIG values are given in Ma.

Historical bins	Sampled fossil record dataset		Recognized fossil record dataset		New discoveries	Cumulative effect of taxonomic work
	Minimum MIG	Maximum MIG	Minimum MIG	Maximum MIG		
1910	2465.4	2573.5	2679.9	2788.9	—	—
1920	2419.7	2528.3	2674.2	2781.8	35.5	−29.1
1930	2363.8	2468.5	2659.7	2767.7	67.9	−53.6
1940	2343.2	2455.0	2640.7	2753.4	17.1	−0.4
1950	2343.2	2455.0	2640.7	2753.4	0.00	0.00
1960	2320.5	2427.4	2598.0	2708.4	25.15	18.7
1970	2229.2	2334.3	2486.9	2599.4	92.2	17.9
1980	1984.3	2091.6	2179.7	2285.9	243.8	66.5
1990	1716.2	1811.0	1856.1	1957.0	274.3	51.9
2000	1485.7	1578.0	1585.0	1676.4	231.8	44.1
2010	1339.0	1423.1	1339.0	1423.1	150.8	98.9

current value of 1.38 Gyr (Table 1). The observed pattern of MIG reduction between 1911 and 1970 is gradual, averaging only 31.9 Myr per decade. Beginning with the 1980 historical bin, the rate of global MIG reduction drastically increases, resulting in a distinct deflection in the slope of the graphed data in Figure 3. After this inflection point (i.e., from 1971 to 2010), global MIG reduction averages 290.5 Myr per decade. The coefficient of determination (r^2 value) for the linear regression line fit to the 1910–1970 data is 0.747. Conversely, the r^2 value for the linear regression line for post-1970 portion of the graph is 0.997.

Figure 4 displays the differential pattern of reduction in global MIG values for both data sets, with the Recognized Fossil Record data set values shown in black. Peak reduction in observed global MIG values occurs between 1981 and 1990. During the following two historical bins, the rate of global MIG reduction decreases, reaching only ∼77% the 1980 level in the 2010 historical bin.

Differences between the Recognized and Sampled Fossil Record Data Sets.—The differences between the values from the Recognized and Sampled Fossil Record data set provide insight into the effects of taxonomic work on our understanding of the avian fossil record over the past century. As expected, the resulting values for the Recognized Fossil Record data set are consistently higher than those obtained

from the Sampled Fossil Record data set (Table 1, Fig. 3), except for in the 2010 historical bin where the values are necessarily equal given the methodology used to calculate global MIG (see Fig. 2). Beginning in the 1910 historical bin, the difference between the average values for each historical bin in both data sets is 215.0 Myr. Thus, roughly 8% of the 1910 observed missing fossil record resulted from the inaccurate referral of known fossils. From
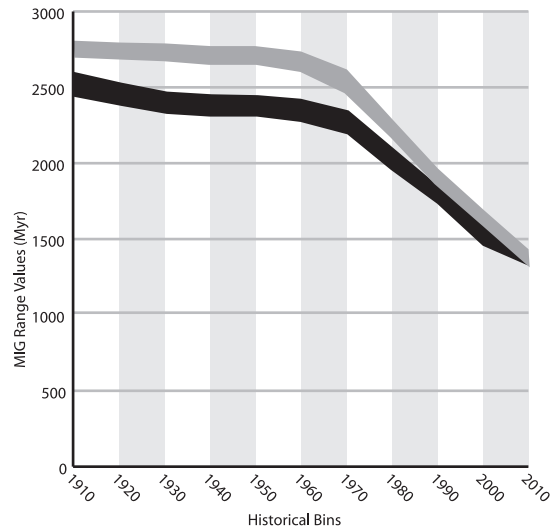

FIGURE 3.   Global MIG values calculated for 75 major lineages within Aves over historical time using the Recognized Fossil Record data set (gray band) and the Sampled Fossil Record data set (black band). See Table 1 for values. Note inflection point at the transition between the pre-1970 and post-1970 periods. Values for each data set are equal in the 2010 historical bin.
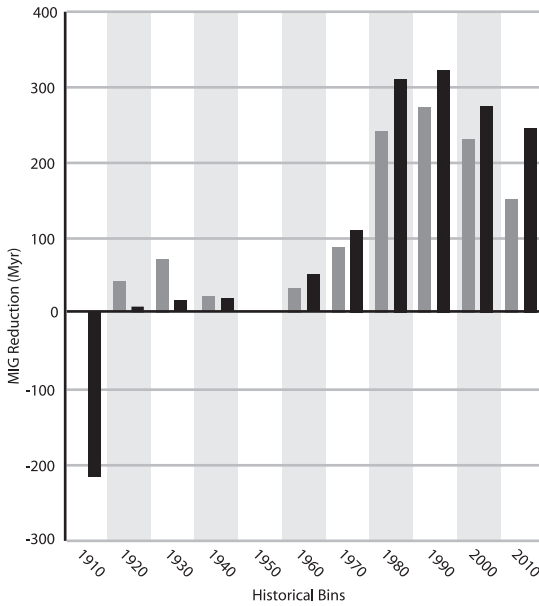
FIGURE 4. Graph of the predicted (gray bars) and observed (black bars) decrease in global MIG values for each historical bin. Predicted values are based on the Sampled Fossil Record data set, which takes into account all fossils discovered by the end date of each historical bin. Observed values are based on the Recognized Fossil Record data set, which excludes fossils that had been discovered but were not yet recognized as members of the relevant clade by the end date of each bin. The initial negative value for the observed values indicates the disparity between the values of each data set for the 1910 historical bin.

MIG reduction outpaces the rate predicted the by the Sampled Fossil Record data set for the first time. For the remainder of the century the combined effects of taxonomic work and discovery of new fossils outpaces the rate predicted by the Sampled Fossil Record data set (Table 1, Fig. 4). As noted above, the highest rate of observed reduction in global MIG values occurred during the 1990 historical bin and the highest predicted rate of global MIG reduction based solely on the discovery of new fossils occurs in the 1990 historical bin. Immediately following this peak, the predicted rate of global MIG reduction falls by ~45% between 1990 and 2010.

Discussion

The Improving Fossil Record of Birds.—Major reviews of the avian fossil record often begin by noting how our knowledge of fossil birds has increased markedly since the last compilation (Lambrecht 1933; Brodkorb 1963; Olson 1985; Mayr 2009b). Our results show a dramatic reduction in total MIG for Aves over the last century: for the Recognized Fossil Record data set global MIG values decline from 2.73 Gyr in 1910 to 1.38 Gyr today (Table 1). In this sense, it can be said that the ghost lineages in the avian fossil record have been cut roughly in half over the course of the past 100 years.

A distinct change in slope of the best-fit linear trend line is evident between the 1910–1960 and 1970–2010 portions of both the Sampled and the Recognized Fossil Record data set curves (Fig. 3). This inflection represents a major turning point in the collection, description, and systematic revision of fossil birds, corresponding to an order of magnitude increase in rate of discovery (31.9 Myr to 290.5 Myr for the Recognized Fossil Record data set). Over the 1910–1960 interval, global MIG reduction for the Recognized Fossil Record data set averages only 31.9 Myr per decade (or ~1% of the 1910 MIG level), corresponding to a trivial improvement in the overall fossil record of the major lineages. After the inflection point (1970–2010), global MIG reduction averages 290.5 Myr per decade (or ~11% of the 1910 MIG level). Although several prolific Lagerstätten such as

1920 through 1950, the average difference between each data set increases or remains constant (see Table 1: Cumulative effect of taxonomic work). In the 1960 and 1970 historical bins the average difference decreases slightly (18.7 Myr and 17.9 Myr, respectively). However, beginning with the 1980 historical bin the average differences between the values from each data set decrease substantially, continuing on until the present day.

In Figure 4, the values for the Sampled Fossil Record data set (gray columns) indicate how much reduction in global MIG should have occurred during each historical bin due to new discoveries alone, assuming all taxonomic referrals of new material were accurate (see also Table 1: New discoveries). Alternatively, the values for the Recognized Fossil Record data set (Fig. 4: black columns) show the observed pattern of global MIG reduction. From 1951 to 1960, the observed rate of global

the Green River Formation and Messel Formation were heavily collected during the 1970–2010 interval, a close inspection of the primary data shows that no one locality contributes disproportionately to this rate increase. Messel and Green River fossils collected after 1970 account for the oldest records of only four and five of the 75 sampled lineages, respectively. With the exceptions of *Prefica nivea* (Steatornithidae) and *Limnofregata azygosternon* (Fregatidae), which eliminate ~50 Myr of MIG from the previously unsampled ranges of their respective clades, fossils from the Messel Formation and Green River Formation result in modest MIG reductions compared to the next geologically oldest records. A total of 40 localities, many divisible into sublocalities or horizons (e.g., the Quercy Fissure Fills) contribute to MIG reduction during the 1970–2010 interval. Many of these localities were actively quarried before 1970, often with an emphasis on non-avian fossils. This pattern suggests that an increase in avian paleontologist worker effort is a more influential driver of the inflection than discovery of a few key fossil sites. Thus, this inflection may be a result of paleornithologists beginning to tackle the glut of undescribed avian fossils that were previously languishing in existing museum collections.

Effects of Discovery versus Revisions.—Notably, the change in slope at the 1910–1960/1970–2010 transition is more pronounced in the Recognized Fossil Record data set than in the Sampled Fossil Record data set. Over the 1910–1960 interval, slightly more MIG is filled in during each decade in the Sampled Fossil Record data set than in the Recognized Fossil Record data set (39.6 Myr versus 31.9 Myr average MIG reduction per decade). This indicates that relevant fossils were being discovered, but were also regularly being misidentified during this historical time period. In contrast, during the 1970–2010 interval an average of 225.2 Myr of MIG was filled in per decade in the Sampled Fossil Record data, but 290.5 Myr of MIG was filled in per decade in the Recognized Fossil Record data. The difference (65.3 Myr average per decade) is due to taxonomic revision of previously described fossils. We hypothesize

that these large gains correspond to the advent of cladistic methodologies and their adoption by the avian paleontology community. Although few actual phylogenetic analyses including fossils were conducted in the 1960–1980 interval, synapomorphy-based approaches to identifying fossil specimens began to supplant criteria based on simple similarity during this time.

Distribution of Gaps.—Broken down into average MIG per branch, our results indicate a decrease from 18.5 Myr per branch in 1910 to 9.3 Myr per branch in 2010 for the Recognized Fossil Record data set, and from 17.0 Myr in 1910 to 9.3 Myr in 2010 per branch for the Sampled Fossil Record data set. However, discussing the results in terms of average branch length implies that MIG is evenly distributed throughout the tree. On the contrary, several lineages have witnessed major reductions in MIG over the past few decades (e.g., Anseranatidae, Steatornithidae, Podargidae, and Passeriformes) whereas other lineages still exhibit very large MIGs (exceeding 40 Myr in Anhimidae, Caprimulgidae, Hemiprocnidae, Opisthocomidae, Psophiidae, and Meropidae). Notably, the vast majority of implied missing fossil record is concentrated on terminal branches, with only a few internodes displaying long ghost lineages (Fig. 5: dashed lines). This observed distribution of MIG is good news, because it implies that much of the currently hypothesized missing avian fossil record will be filled by stem taxa of well-circumscribed extant clades corresponding to traditional orders and families, rather than deeply diverging taxa that belong on lineages leading to large, heterogeneous clades uniting multiple ''orders.'' If this pattern reflects reality, accurate systematic placement of fossils filling the major remaining gaps should be relatively easier when and if they are found.

Future Trends.—One of the most pertinent questions for active paleornithologists that this study addresses is what predictions these data make for future discoveries of oldest known records for avian clades. Given the patterns observed here, are reductions in global MIG values expected to continue at the observed rate, or do the data suggest we are approaching
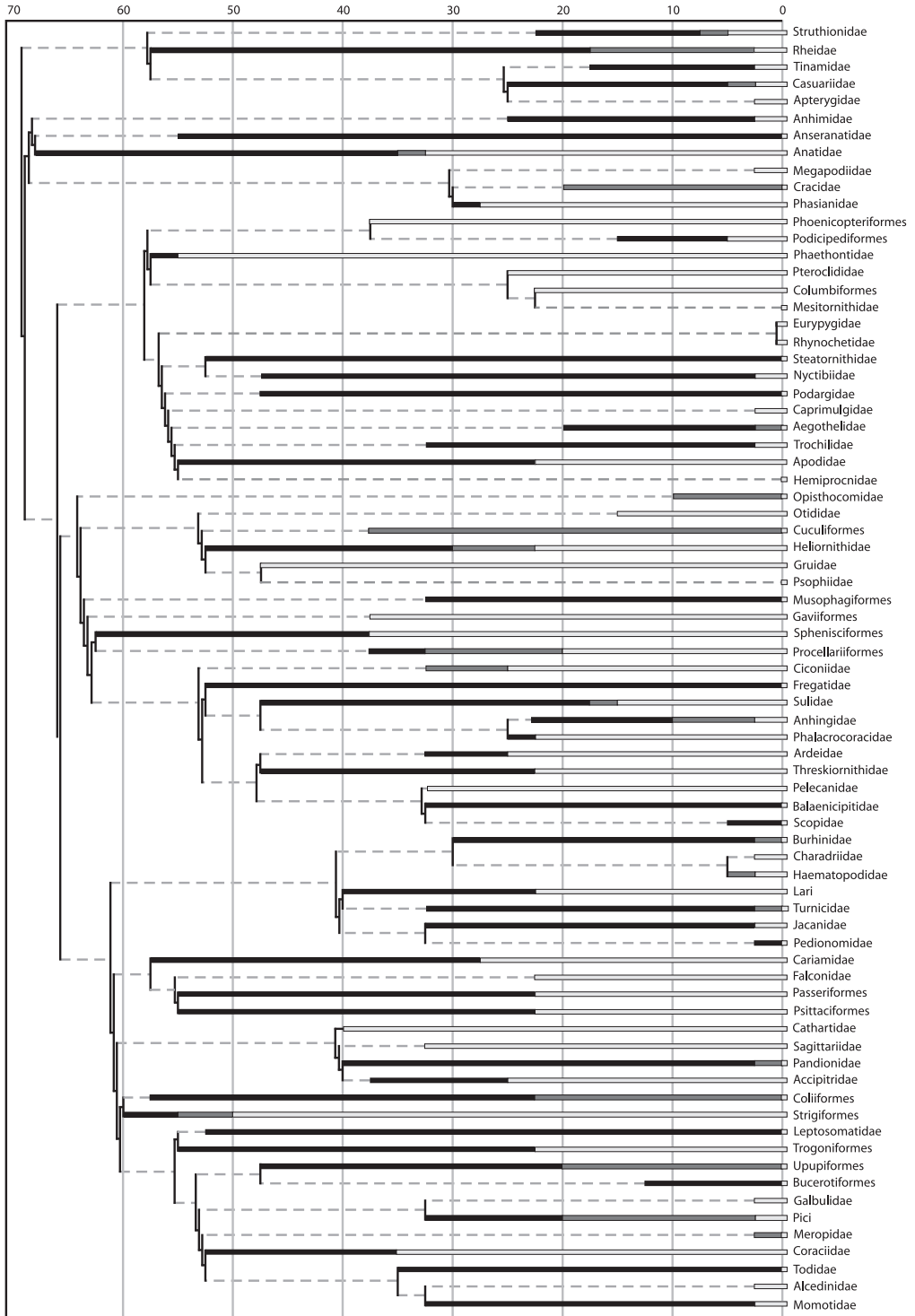
FIGURE 5. Time-calibrated cladogram generated using the phylogeny of Hackett et al. (2008) and the Recognized Fossil Record data set illustrating the temporal ranges for the 75 avian clades examined in this study. White bars indicate the fossil record as known in 1910, gray bars indicate temporal range extensions that occurred between 1911 and 1970, and black bars indicate temporal range extensions that occurred between 1971 and 2010. Ghost lineages implied by the 2010 temporal data are indicated by dashed lines.

an asymptote where discoveries affecting global MIG values will become rare? Although these results do not make a clear projection either way, examination of the global MIG data for the 1980 to 2010 historical bins for both data sets provides some interesting insights. Both the Sampled and Recognized Fossil Record data sets show a slowing in the rate of global MIG reduction over the past 20 years. Between these two data sets, the Sampled Fossil Record data set shows the most dramatic decrease in global MIG reduction rate, dropping by ~45% between 1990 and 2010 (Table 1, Fig. 4). However, before making predictions based on that data set, we need to remember the assumptions that were made when constructing that data set. Specifically, these data assume that all present-day taxonomic referrals are correct. Looking over the historical trends (Figs. 3, 4) we can see that this is unlikely to be true. There are almost certainly known fossils in collections that represent unrecognized, oldest records for avian clades. Future identification and taxonomic revision of these fossils will change our current understanding of the predicted rate of global MIG reduction, reducing the disparity between the 2000 and 2010 historical bins for the Sampled Fossil Record data set. However, the pattern displayed by the values from the Recognized Fossil Record data set will not be affected by future taxonomic revisions or discoveries, because their calculation is based on historical understandings of the fossil record that are not subject to future revision. Therefore, the observed trend of decreasing rate of global MIG reduction in the Recognized Fossil Record data set is real. What remains to be determined is whether this decline represents a temporary lull, or the beginning of a pronounced decline in the rate of global MIG reduction that extends into the foreseeable future.

A recurring debate in avian systematics is the source of the extreme discord between molecular divergence dating estimates for the age of the extant radiation of birds and the age provided by the fossil record (Cooper and Penny 1997; Cooper and Fortey 1998; Cracraft 2001; Ericson et al. 2006; Clarke et al. 2007; Ksepka and Clarke 2010). Sequence-based divergence studies have almost universally supported divergences for many extant avian

clades deep in the Cretaceous (e.g., Barker et al. 2004; Brown et al. 2008). However, only a few basal clades can be inferred as present in the Late Cretaceous from the fossil record and phylogeny (Clarke et al. 2005), and a cryptic Early Cretaceous radiation for crown Aves seems unlikely given the extensive fossil record of non-crown birds (Feduccia 1999; Fountaine et al. 2005; Mayr 2009b). Given the substantial rate of improvement in the fossil record documented here for the past few decades, it is worth examining whether the fossil record may be "catching up" to molecular divergence estimates. Although major reductions in MIG have occurred throughout the avian tree over the past 100 years, the gaps filled fall almost entirely in the Cenozoic. Despite the rapid rate of discovery of gap-filling fossils over the last 100 years (Fig. 5), the fossil record of crown Aves as a whole was pushed back by only about 12 Myr, from the early Eocene into the Late Cretaceous. Thus the answer to this question is mixed: Historical trends suggest that new discoveries or revisions are likely to bring the fossil records of many clades a few million years closer to molecular divergence estimates. However, it remains unlikely that the fossil record of crown Aves will be pushed back into the Early Cretaceous, especially if the declining rate of new discoveries noted in the 2010 time bin represents a long-term trend. Another issue to consider is whether concordance represents a moving target. Because the oldest fossils are highly sought out as calibration points for divergence analyses, new discoveries could continue pushing back molecular divergence dates depending on the methods applied. There are known biases toward overestimating the length of deep internal branches in many divergence dating methods (Phillips 2009), and unless these are addressed, discordance will likely continue regardless of new fossil discoveries. One interesting course of investigation beyond the scope of this study would be to use the historical data sets compiled here to estimate molecular sequence-based divergence estimates for each time bin (using the fossil record as known at that time) to test how new discoveries would have changed the predicted divergence dates in a historical

context. This could address whether divergence dates calculated using the present-day fossil record to provide calibrations are substantially older than those derived using the fossil record available in 1910, or would remain relatively stable.

On a final note, our results indicate that despite the recent decline in the rate of global MIG reduction, there is still a vast amount of progress to be made before we can consider the fossil record of birds to be ''completely'' sampled. Between 1.34 and 1.42 billion years total MIG still remains between the predicted and actual first appearances of the major crown lineages, assuming that the oldest known record of crown Aves does not change. As discussed above, even this imposing value is a minimum because it takes the 2010 discovery data as a baseline and new discoveries may either add or subtract global MIG depending on where they fall in the tree. Sequence-based divergence studies have in some cases supported divergences for many extant avian clades deep in the Cretaceous (e.g., Barker et al. 2004; Crowe et al. 2006; Baker et al. 2007; Brown et al. 2008), which would require an order of magnitude level increase in MIG. However, historical patterns of improving completeness in the fossil record do not support the assumption that new discoveries will push the OKRs of multiple crown bird lineages deep into the Cretaceous (Fara and Benton, 2000; this study). Nonetheless, even a more plausible Late Cretaceous radiation for most major subclades of Aves would require approximately doubling the inferred amount of global MIG. Regardless of whether the age of the basal split in Aves occurred in the Early Cretaceous or Late Cretaceous, avian paleontologists clearly have much work to do both in the field and in the lab if the vast expanses of missing record are to be even partially filled.

The results presented here provide the first quantitative analysis of the factors and trends influencing our understanding of the avian fossil record. Given the unique history of avian paleontology, the patterns uncovered here do not necessarily represent general trends across vertebrates. Future comparisons with other clades such as mammals or turtles will reveal whether similar trends in the Sampled Fossil Record and Recognized Fossil Record occur in other groups, and allow us to determine whether clades favored with ''better'' fossil records actually show comparable amounts of global MIG at equivalent taxonomic resolutions.

Acknowledgments

Literature Cited

Andors, A. 1992. Reappraisal of the Eocene groundbird *Diatryma* (Aves: Anserimorphae). Natural History Museum of Los Angeles County, Science Series 36:109–125.

Angielczyk, K. D., and A. A. Kurkin. 2003. Phylogenetic analysis of Russian Permian dicynodonts (Therapsida: Anomodontia): implications for Permian biostratigraphy and Pangaean biogeography. Zoological Journal of the Linnean Society 139:157–212.

Baker, A. J., S. L. Pereira, and T. A. Paton. 2007. Phylogenetic relationships and divergence times of Charadriiformes genera: multigene evidence for the Cretaceous origin of at least 14 clades of shorebirds. Biology Letters 3:205–209.

Bardack, D. 1991. First fossil hagfish (Myxinoidea): a record from the Pennsylvanian of Illinois. Science 254:701–703.

Bardack, D., and R. Zangerl. 1968. First fossil lamprey: a record from the Pennsylvanian of Illinois. Science 162:1265–1267.

Barker, F. K., A. Cibois, P. Schikler, J. Feinstein, and J. Cracraft. 2004. Phylogeny and diversification of the largest avian radiation. Proceedings of the National Academy of Sciences USA 101:11040–11045.

Benton, M. J. 1994. Paleontological data and identifying mass extinctions. Trends in Ecology and Evolution 9:181–185.

———. 1999. Early origin of modern birds and mammals: molecules versus morphology. BioEssays 21:1043–1051.

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

Benton, M. J., M. Wills, and R. Hitchin. 2000. Quality of the fossil record through time. Nature 403:534–537.

Bleiweiss, R. 1998. Fossil gap analysis supports early Tertiary origin of trophically diverse avian orders. Geology 26:323–326.

Bourdon, E. 2005. Osteological evidence for sister group relationship between pseudo-toothed birds (Aves: Odontopterygiformes) and waterfowls (Anseriformes). Naturwissenschaften 92:586–591.

Boyd, C. A., T. P. Cleland, N. L. Marrero, and J. A. Clarke. 2011. Exploring the effects of phylogenetic uncertainty and consensus trees on stratigraphic consistency scores: a new program and a standardized method. Cladistics 27:52–60.

Brodkorb, P. 1963. Catalog of fossil birds, Part 1. Archaeopterygiformes through Ardeiformes. Bulletin of the Florida State Museum Biological Sciences 7:179–293.

Brown, J., J. Rest, J. Garcia-Moreno, M. Sorenson, and D. Mindell. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biology 6:6.

Clarke, J. A., C. P. Tambussi, J. I. Noriega, G. M. Erickson, and R. A. Ketcham. 2005. Definitive fossil evidence for the extant avian radiation in the Cretaceous. Nature 433:305–308.

Clarke, J. A., D. T. Ksepka, M. Stucchi, M. Urbina, N. Giannini, S. Bertelli, Y. Narvaez, and C. A. Boyd. 2007. Paleogene equatorial penguins challenge the proposed relationship between biogeography, diversity, and Cenozoic climate change. Proceedings of the National Academy of Sciences U.S.A. 104:11545–11550.

Cooper, A., and D. Penny. 1997. Mass Survival of birds across the Cretaceous-Tertiary boundary: molecular evidence. Science 275:1109–1113.

Cooper, A. and R. Fortey. 1998. Evolutionary explosions and the phylogenetic fuse. Trends in Ecology and Evolution 13:151–156.

Cracraft, J. 1980. Phylogenetic theory and methodology in avian paleontology: a critical appraisal. Contributions in Science of the Natural History Museum of Los Angeles County 330:9–16.

———. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. Proceedings of the Royal Society of London B 268:459–469.

Cracraft, J., F. K. Barker, J. Braun, J. Harshman, G. J. Dyke, J. Feinstein, S. Stanley, A. Cibois, P. Schikler, P. Beresford, J. García-Moreno, M. D. Sorenson, T. Yuri, and D. P. Mindell. 2004. Phylogenetic relationships among modern birds (Neornithes): towards an avian tree of life. Pp. 468–489 in J. Cracraft and M. J. Donoghue, eds. Assembling the tree of life. Oxford Press, New York.

Crowe, T. M., R. C. K. Bowie, P. Bloomer, T. G. Mandiwana, T. A. J. Hedderson, E. Randi, S. L. Pereira, and J. Wakeling. 2006. Phylogenetics, biogeography and classification of, and character evolution in, gamebirds (Aves: Galliformes): effects of character exclusion, data partitioning and missing data. Cladistics 22:495–532.

Ericson, P. G. P., C. L. Anderson, T. Britton, A. Elzanowski, U. S. Johansson, M. Källersjö, J. I. Ohlson, T. J. Parsons, D. Zuccon, and G. Mayr. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biology Letters 4:543–547.

Fara, E., and M. J. Benton. 2000. The fossil record of Cretaceous tetrapods. Palaios 15:161–165.

Feduccia, A. 1999. The origin and evolution of birds. Yale University Press, New Haven, Conn.

Foote, M., and J. J. Sepkoski Jr. 1999. Absolute measures of the completeness of the fossil record. Nature 389:415–417.

Fountaine, T. M. R., M. J. Benton, G. J. Dyke, and R. L. Nudds. 2005. The quality of the fossil record of Mesozoic birds. Proceedings of the Royal Society of London B 272:289–294.

Gauthier, J., A. G. Kluge, and T. Rowe. 1988. Amniote phylogeny and the importance of fossils. Cladistics 4:105–209.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K.-L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt, and T. Yuri. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763–1768.

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

Jeffery, C. H., and R. B. Emlet. 2003. Macroevolutionary consequences of developmental mode in temnopleurid echinoids from the Tertiary of southern Australia. Evolution 57:1031–1048.

Kalmar, A., and D. J. Currie. 2010. The completeness of the continental fossil record and its impact on patterns of diversification. Paleobiology 36:51–60.

Kerr, A. M., and J. Kim. 2001. Phylogeny of Holothuroidea (Echinodermata) inferred from morphology. Zoological Journal of the Linnean Society 133:63–81.

Ksepka, D. T., and J. A. Clarke. 2010. The basal penguin (Aves, Sphenisciformes) *Perudyptes devriesi* and a phylogenetic evaluation of the penguin fossil record. Bulletin of the American Museum of Natural History 337:1–77.

Lambrecht, K. 1933. Handbuch der Palaeornithologie. Gebrüder Bornträger, Berlin.

Marjanovic, D., and M. Laurin. 2007. Fossils, molecules, divergence times, and the origin of lissamphibians. Systematic Biology 56:369–388.

Maxwell, W. D., and M. J. Benton. 1990. Historical tests of the absolute completeness of the fossil record of tetrapods. Paleobiology 16:322–335.

Mayr, G. 2002. Osteological evidence for paraphyly of the avian order Caprimulgiformes (nightjars and allies). Journal of Ornithology 143:82–97.

———. 2005. Tertiary plotopterids and a novel hypothesis on the phylogenetic relationships of penguins (Spheniscidae). Journal of Zoological Systematics and Evolutionary Research 43:63–71.

———. 2007. The renaissance of avian paleontology and its bearing on the higher-level phylogeny of birds. Journal of Ornithology 148(Suppl. to No. 2):S455–S548.

———. 2009a. Phylogenetic relationships of the paraphyletic 'caprimulgiform' birds (nightjars and allies). Journal of Zoological Systematics and Evolutionary Research.

———. 2009b. Paleogene fossil birds. Springer, Heidelberg.

Murray, P. F., and O. Vickers-Rich. 2004. Magnificent mihirungs: the colossal flightless birds of the Australian dreamtime. Indiana University Press, Bloomington.

Newell, N. D. 1959. Adequacy of the fossil record. Journal of Paleontology 33:488–499.

Norell, M. A. 1992. Taxic origin and temporal diversity: the effect of phylogeny. Pp. 88–118 in M. J. Novacek and Q. D. Wheeler, eds. Extinction and phylogeny. Columbia University Press, New York.

Norell, M. A., and M. J. Novacek. 1992. The fossil record and evolution: comparing cladistic and paleontologic evidence for vertebrate history. Science 255:1690–1693.

Olson, S. L. 1976. Oligocene fossils bearing on the origins of the Todidae and the Momotidae (Aves: Coraciiformes). *In* S. L. Olson, ed. Collected papers in avian paleontology honoring the 90th birthday of Alexander Wetmore. Smithsonian Contributions to Paleobiology 27:111–119.

———. 1985. The fossil record of birds. Pp. 79–238 in D. S. Farner, J. R. King, and K. C. Parkes, eds. Avian biology. Academic Press, New York.

Olson, S. L., and Y. Hasegawa. 1979. Fossil counterparts of giant penguins from the North Pacific. Science 206:688–689.

———. 1996. A new genus and two new species of gigantic Plotopteridae from Japan (Aves: Pelecaniformes). Journal of Vertebrate Paleontology 16:742–751.

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

Phillips, M. J. 2009. Branch-length estimation bias misleads molecular dating for a vertebrate mitochondrial phylogeny. Gene 441:132–140.

Pol, D., and M. A. Norell. 2006. Uncertainty in the age of fossils and the stratigraphic fit to phylogenies. Systematic Biology 55:512–521.

Pol, D., M. A. Norell, and M. E. Siddall. 2004. Measures of stratigraphic fit to phylogeny and their sensitivity to tree size, shape and scale. Cladistics 20:64–75.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. Science 177:1065–1071.

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

Stresemann, E. 1959. The status of avian systematics and its unsolved problems. Auk 76:269–289.

Valentine, J. W. 1970. How many marine invertebrate fossils? Journal of Paleontology 44:410–415.

von Meyer, H. 1839. Ein vogel im Kreideschiefer des Kantons *Glaris*. Neues Jahrbuch für Mineralogie, Geognosie, Geologie und Petrefaktenkunde 1:683–685.

———. 1844. [Letter]. Neues Jahrbuch für Mineralogie, Geognosie, Geologie und Petrefaktenkunde 6:329–340.

Walsh, S. L. 1998. Fossil datum and paleobiological event terms, paleontostratigraphy, chronostratigraphy, and the definition of land mammal ''age'' boundaries. Journal of Vertebrate Paleontology 18:150–179.

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

Wills, M., P. M. Barrett, and J. F. Heathcote. 2008. The modified gap excess ratio (GER*) and the stratigraphic congruence of dinosaur phylogenies. Systematic Biology 57:891–904.