Biological Journal of the Linnean Society, 2024, **XX**, 1–11
https://doi.org/10.1093/biolinnean/blae070
Advance access publication 23 August 2024
Original Article

Original Article

An evaluation of the usefulness of morphological characters to infer higher-level relationships in birds by mapping them to a molecular phylogeny

Per G.P. Ericson FLS[1,*] and Yanhua Qu[1,2,*]

[1]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, PO Box 50007, SE-104 05, Stockholm, Sweden
[2]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, 100101, China

[*]Corresponding authors. Department of Bioinformatics and Genetics, Swedish Museum of Natural History, PO Box 50007, SE-104 05, Stockholm, Sweden. E-mail: per.ericson@nrm.se; Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, 100101, China. E-mail: quyh@ioz.ac.cn

ABSTRACT

The use of genetic data to reconstruct systematic relationships has revolutionized our understanding of avian evolution. Morphology-based classifications were often in conflict because of different opinions among scholars about the relative importance of certain phenotypes. The considerable morphological variation observed among birds was codified into phylogenetic characters by Livezey and Zusi (2006) who also scored them for 150 extinct and extant taxa. Herein we have evaluated the phylogenetic signal of 1860 of these characters by mapping them to a molecular phylogeny including 102 taxa that represent all extant birds (with the underlying assumption that this tree topology is a good estimate of the evolutionary relationships among birds). The characters fit the molecular tree with a mean consistency index (CI) of 0.38. Muscle characters are the most homoplasious (CI 0.32), while characters related to integument, feathers, intestinal, respiratory, syrinx, urogenital, nervous, and reproductive organs show a considerably better fit (mean CI 0.49). We also explored what characters may unambiguously support certain basal clades that are well-supported by molecular data. We found only a few clades (e.g. Galloanserae, Procellariimorphae) being supported by unambiguous apomorphies, while many well-established clades (e.g. Pelecaniformes, Charadriiformes, Accipitriformes, Coraciiformes) lack such support entirely.

Keywords: morphology; convergent evolution; homoplasy; birds; higher-level systematics

INTRODUCTION

Early classifications of birds into 'natural' groups long relied on differences in their external anatomy, and characters related to their style of feeding or locomotion were considered particularly important. In the 19th century, these 'beak-and-feet' classifications were gradually replaced by studies of various aspects of the internal morphology, and considerable attention was given to, for example, osteology, myology, tendons, and intestines. Impressively ambitious studies by Max Fürbringer, Alfred Henry Garrod, Hans Gadow, and several others resulted in a wealth of anatomical information, but also in a growing realization that the morphology, to a large degree, carried contradictory information about systematic relationships. This sparked considerable debate about what kind of characters provide the most reliable evidence for a shared ancestry between groups of birds. Towards the end of the 19th century, along with the growing acceptance of evolutionary theory, these discussions often focused on how to distinguish advanced ('derived') characters from primitive (reptile-like) ones. Based upon individual judgments by the systematists, several different hierarchical classifications were proposed to best reflect the evolution of birds. A major concern for all systematists was that the large amount of conflicting data could not easily be translated into logical arrangements using pen and paper (this of course became easier with the development of computers during the last decades of the 20th century). At that point in time the confidence in the different classifications relied heavily on the authority of the scholar who had proposed it.

Of particular importance to the use of morphological data in systematics is the scientific battle concerning philosophical principles and analytical methods that began with a publication by Willi Hennig (1950). According to Hennig, all relationships between taxa must be based on derived (homologous) characters,

Received 24 April 2024; revised 4 June 2024; accepted 12 July 2024

since convergently evolved (analogous) characters bear no information relevant for systematic inference. While the importance of derived characters had long been realized, Hennig (1950) developed the theory behind it and formalized the principles of the new methodology, which was named 'cladistics'. However, the cladistic method did not solve the problem of conflicting characters in practice; the problem of separating homologous characters from analogous remained. A practical solution was offered by the introduction of the principle of maximum parsimony, i.e. the idea that the most probable phylogenetic hypothesis is the one that requires the smallest number of *ad hoc* explanations (Kluge and Farris 1969, Farris 1970, Fitch 1971). In reality, this meant that one should count the number of times each character has changed from one state to another in the tree and sum up this for all characters. The phylogeny that requires the smallest total number of character state transformations provides the most parsimonious explanation of the observed data. The idea of using maximum parsimony to discriminate between alternative phylogenetic hypotheses stirred up enormous scientific debate over several decades (cf. Hull 1988). In ornithology, the debate was fierce and compounded by early studies that aimed at demonstrating the usefulness of cladistics methods in avian systematics utilized poorly chosen characters, leading to spurious relationships that were easy to ridicule (cf. Olson 1982). What plagued these studies were questionable assumptions about the homology of both the included morphological characters and the different character states identified. To some researchers the answer to the dilemma of uncertain homologies was to add more characters. The underlying assumption was that the information from the non-homologous (homoplastic) characters would produce a random noise and that the phylogenetic signal would be discernible if only enough characters were included in the analysis (Kluge 1989, see also Rieppel 2005). A benefit of this approach is that it does not require ensuring the homology of each character by the use of embryology or other (tedious) methods.

The use of molecular sequence data for phylogenetic inference has been dominating phylogenetic research since the mid-1990s, when researchers could abandon the laborious methods based of chromatography to obtain DNA sequences, and instead generate these data using PCR amplification and Sanger sequencing. In the early days of DNA-based systematics, the sequence data was produced mostly by people more knowledgeable about genetic methods than organismal systematics. At that time, most molecular datasets consisted of only one or a few mitochondrial markers. Even then, there were complaints among morphological systematists that the molecular datasets would 'swamp' most morphological datasets available for the same group of taxa. It felt genuinely unfair that years of work to identify homologous and systematically meaningful morphological characters resulted in a dataset that was many times smaller than the potentially informative nucleotides in an alignment of a single mitochondrial gene. The sentiment especially among morphologists was then that one should produce separate phylogenetic trees from the molecular and morphological data, respectively, and then search for congruence between them (Nixon and Carpenter 1996). Later on, it became more common to combine the two categories of data into a single dataset (Gatesy and Arctander 2000, Murrell *et al.* 2001, Gatesy *et al.* 2003), especially after several studies had shown that morphological datasets may contain more phylogenetic signal per character and thus prevail also in much larger molecular datasets (e.g. Baker *et al.* 1998). The latter observations highlighted the fact that homoplasy might be a considerable problem also in certain molecular datasets, especially when using mitochondrial data to infer relationships of taxa that split many tens of million years ago. However, as more and more nuclear markers were used for phylogenetic inference there was a growing feeling that these different molecular datasets produced largely congruent results (Wortley and Scotland 2006). Today, phylogenetic hypotheses in ornithology rely almost entirely on molecular data and there is a wide consensus about the major patterns of avian relationships.

In the late 1990s, when molecular datasets used to estimate bird relationships still mostly consisted of only a few mitochondrial and nuclear markers, it became clear that these would not be able to resolve the avian tree of life. There was also a feeling that the vast information from morphology and behavioural variation was poorly utilized in systematics and if only evaluated in a consistent manner and analysed with cladistic methods, these data had potential to increase the understanding of avian relationships. This led avian anatomists Richard L. Zusi and Bradley C. Livezey to undertake a multi-year, morphologically-based phylogenetic study of modern birds (Neornithes), which eventually also received financial support from the National Science Foundation in the USA (Livezey and Zusi 2001). The rationale for this was that they 'believed that comparative anatomy deserves a multisystemic, uniformly characterized, and taxonomically inclusive assessment to provide balance to the enthusiasm and widespread resources being bestowed upon molecular techniques' (Livezey and Zusi 2006: p.2). Their work resulted in two major publications, one that constituted a detailed description of the 2954 morphological (2451 osteological, 256 myological, and 247 miscellaneous) characters identified by the authors (Livezey and Zusi 2006), and other in which these characters were analysed phylogenetically (Livezey and Zusi 2007). Together, these publications constitute a major contribution to the knowledge about primarily morphological variation in birds. The importance of this work is vast, but its impact on the understanding of the higher-level relationships (orders and families) in birds did not become as prominent as hoped for (e.g. Mayr 2008). The main reason is that the study was published at a time when nuclear DNA had begun to produce novel and often highly unexpected phylogenetic hypotheses about modern birds. Based on the distribution of insertions in an intron of the nuclear beta-fibrinogen gene, Fain and Houde (2004) suggested a deep split within Neoaves, although the most other parts of the phylogeny were left unresolved. Ericson *et al.* (2006) augmented the dataset with four other nuclear genes and could present the first well-supported molecular phylogeny of Neoaves. Many relationships in that study were highly unexpected, but the results were later corroborated by analyses of increasingly more data (Hackett *et al.* 2008, McCormack *et al.* 2013, Jarvis *et al.* 2014, Prum *et al.* 2015). Today, the major patterns of higher-level relationships in birds are considered as well resolved. The fact that many of the relationships suggested based on the morphological data (Livezey and Zusi 2007) differ dramatically from those based on a wealth of molecular information, has strengthened the opinion that the morphological dataset of Livezey and Zusi suffers from a high degree of homoplasy. Nevertheless, this dataset surely consists also of many characters for which the homology assumptions are correct, and one way to identify these is to map the characters to a tree assumed to be an appropriate representation of our current understanding of the phylogenetic relationships among birds. By this, we may be able to evaluate not only the proposed homology of individual characters and character states, but also to investigate if any part of the avian morphology is especially prone to convergent evolution. Besides overall assessments, we studied this in the different categories of data (osteological, myological, integumental, etc.), as well as in subsets thereof. Herein we use this strategy, the principle of 'reciprocal illumination' (Hennig 1950, 1966), to analyse how the degree of homoplasy varies among categories of characters and taxonomic groups. It should be mentioned that we herein use the term homoplasy (and non-homologous) to denote similar morphologies that are not inherited from their nearest ancestor of the taxa in which it has been observed. We do not here discriminate between those convergently evolved morphologies that have a common genetic base [Hennig (1966) coined the term 'homoiology' for this] and those that are genetically independent.

Morphology would doubtless contribute important information if the homologies were known. For example, it has been shown that adding morphological characters to a molecular dataset increases the support for most clades in a phylogenetic tree of mammals (Lee and Camens 2009). To evaluate if adding morphological characters to a molecular dataset improves the results also in analyses of higher-level relationships in birds, we re-ran the analysis after adding those of the Livezey and Zusi (2006) morphological characters that had the best fit to the 'Prum tree' to another, independent molecular dataset.

Taxonomic identifications of fossils are based on morphology and it is commonly agreed (e.g. Patterson 1981, Mayr 2017) that hypotheses about the systematic relationships of fossils should be based on apomorphies (derived characters that are shared with other taxa). Only taxonomically properly placed fossils are useful for calibration of evolutionary time-trees etc. (e.g. Magallón 2004, Parham *et al.* 2012), and correct assessments of homologies and distributions across taxa are imperative for this. Thus, the identification by Livezey and Zusi (2006) of 2954 morphological characters represents a major source of information that is potentially useful for this purpose. Suits of apomorphic characters may also serve to define higher-level clades of birds, which is of special interest when assigning fossils to modern groups. Herein we aim to identify such characters for several major taxonomic groups in the molecular tree.

MATERIAL AND METHODS

Morphological characters, phylogenetic hypotheses, and tree statistics

As a molecular-based hypothesis of the evolutionary relationships of birds, we used the tree in Prum *et al.* (2015: fig. 1) based on the analysis of a 300 kb alignment obtained from 259 nuclear loci. This tree was trimmed down to include 102 species representing 99 families for which morphological information is available in Livezey and Zusi (2006). The resulting tree (Fig. 1A) is here called the 'Prum tree'. We then mapped the 2954 characters published by Livezey and Zusi (2006) individually to this phylogenetic tree. For each character, we estimated how well it fits to the 'Prum tree' by calculating the consistency index (CI), retention index (RI), and rescaled consistency index (RC). We calculated these *per* character instead of using the overall ('ensemble') statistics as this is supposed better as a general and comparative method to measure homoplasy in a dataset (Archie 1990). We used PAUP* 4.0a (Swofford 2000) to calculate these statistics and applied the 'exclude Uninf' option to include the 1860 characters that were found to be parsimony-informative among the 102 species. CI is calculated by dividing the minimum number of changes in a certain character (m) by the observed number of changes in the specified tree (s) (Kluge and Farris 1969). The largest number is 1, which means that the character changes fit the tree perfect. RI is an index that measures the fraction of potential synapomorphy retained as synapomorphy on the tree (Farris 1989, Mickevich and Lipscomb 1991). This index is defined as $RI = (h - s)/(h - m)$ where s is the observed number of changes for a certain character on the tree, m is the minimum number of changes in that character, and h is the number of extra steps observed in the tree for the same character, i.e. the homoplasy. An RI of 1 for a character indicates that it is completely consistent on the tree, and an RI of 0 indicates that the character has the maximum possible homoplasy on the tree, with none of its states acting as synapomorphies (Mickevich and Lipscomb 1991). A third statistic calculated is the rescaled consistency index (RC) that is the product of the CI and the RI (Farris 1989). While CI is a measurement of homoplasy, RC measures the proportion of maximum observable homoplasy (Goloboff 2022), and it has the advantage over CI to vary between 0 and 1. CI can never be smaller than m/g, where g denotes the greatest possible value of s (Farris 1989).

To obtain comparable tree statistics also for the morphological tree we pruned the Livezey and Zusi dataset to include the same 102 taxa as the molecular 'Prum tree'. This tree (Fig. 1B) is called the 'Livezey & Zusi tree'. To evaluate the effect of the down-trimming of the dataset on the tree topology we estimated a new morphological tree using only the 1860 characters that are parsimony-informative among the 102 taxa (Supporting information, Fig. S1). Naturally, the overall tree length (i.e. the sum of the minimum numbers of changes over all characters for the given tree topology) became shorter after pruning the original tree, but their topologies agree almost fully. The following differences to the tree in Livezey and Zusi (2007) were observed: (i) Ramphastidae is no longer recovered as monophyletic, (ii) Musophagidae and Opisthocomidae are sisters, (iii) Psittacidae is sister to the Cuculidae, Musophagidae, and Opisthocomidae clade, (iv) there are some rearrangements among the charadriiform families, (v) Threskiornithidae and Scopidae are sisters, as are (vi) Ciconiidae and Phoenicopteridae, (vii) Hydrobatidae and Pelecanoididae, (viii) Diomedeidae and Procellariidae, and (ix) Odontophoridae and Phasianidae (Supporting information, Fig. S1). However, the observed topological differences had only marginal effect on the tree statistics for the two morphological trees (Supplementary Material 1), so we decided to use those calculated for the original morphological tree (pruned to include only 102 taxa) in the analyses.

Defining and comparing subsets of morphological characters

As noted above, we found 1860 characters (1449 osteological, 233 myological, and 178 miscellaneous) to be parsimony-informative for

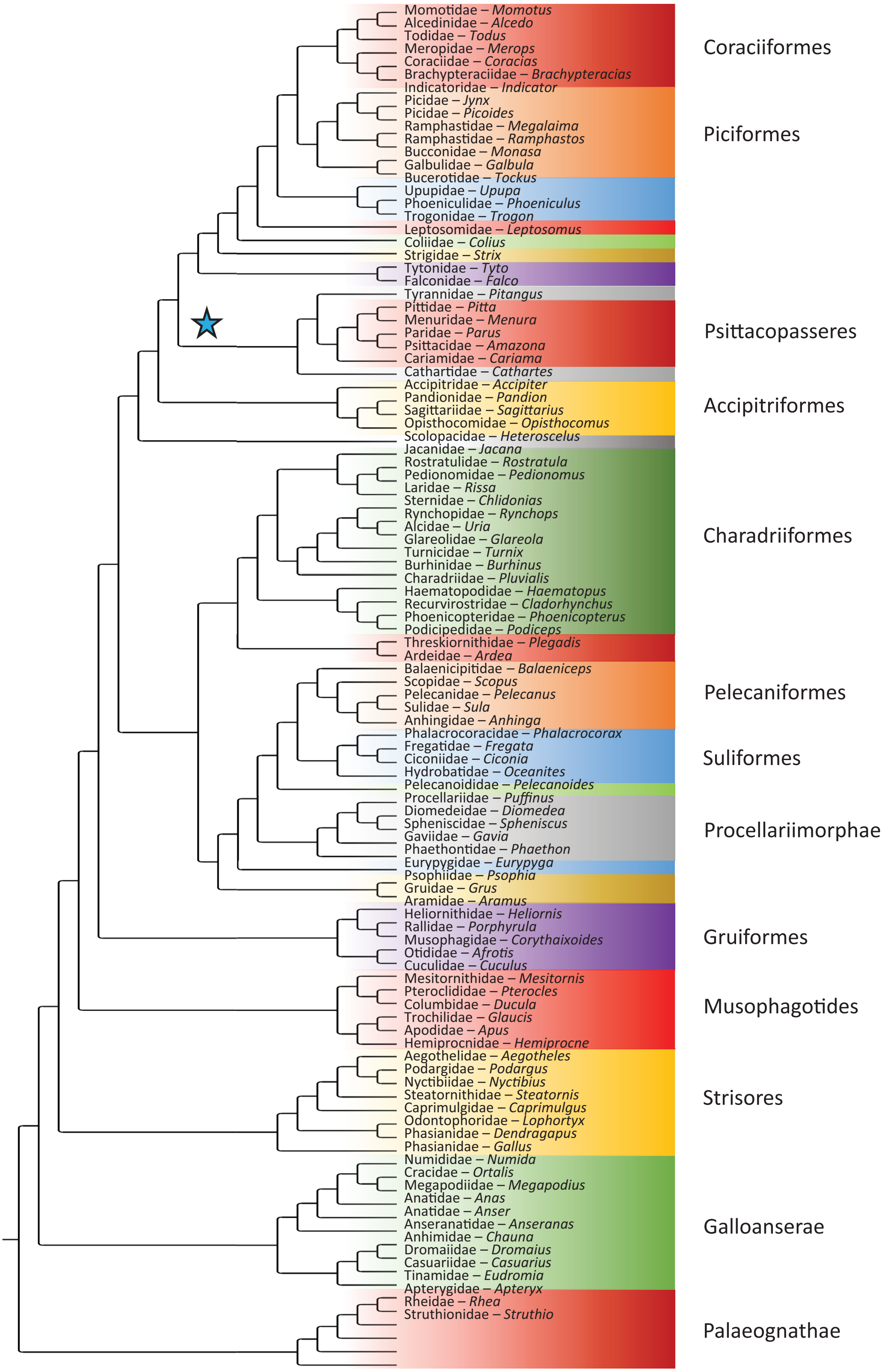

Figure 1. A, the 'Prum tree'. Phylogeny from Prum *et al.* (2015: fig. 1), trimmed down to 102 taxa for which morphological data is available in Livezey and Zusi (2006). B, comparison between topologies of the 'Prum tree' (left) and the 'Livezey & Zusi tree' (right), i.e. a phylogenetic tree based solely on the morphological data in Livezey and Zusi (2007) but pruned to include the same taxa as in the 'Prum tree'. Higher-level clade names follow Sangster *et al.* (2022). The star marks the clade Australavis *sensu* Ericson (2012).

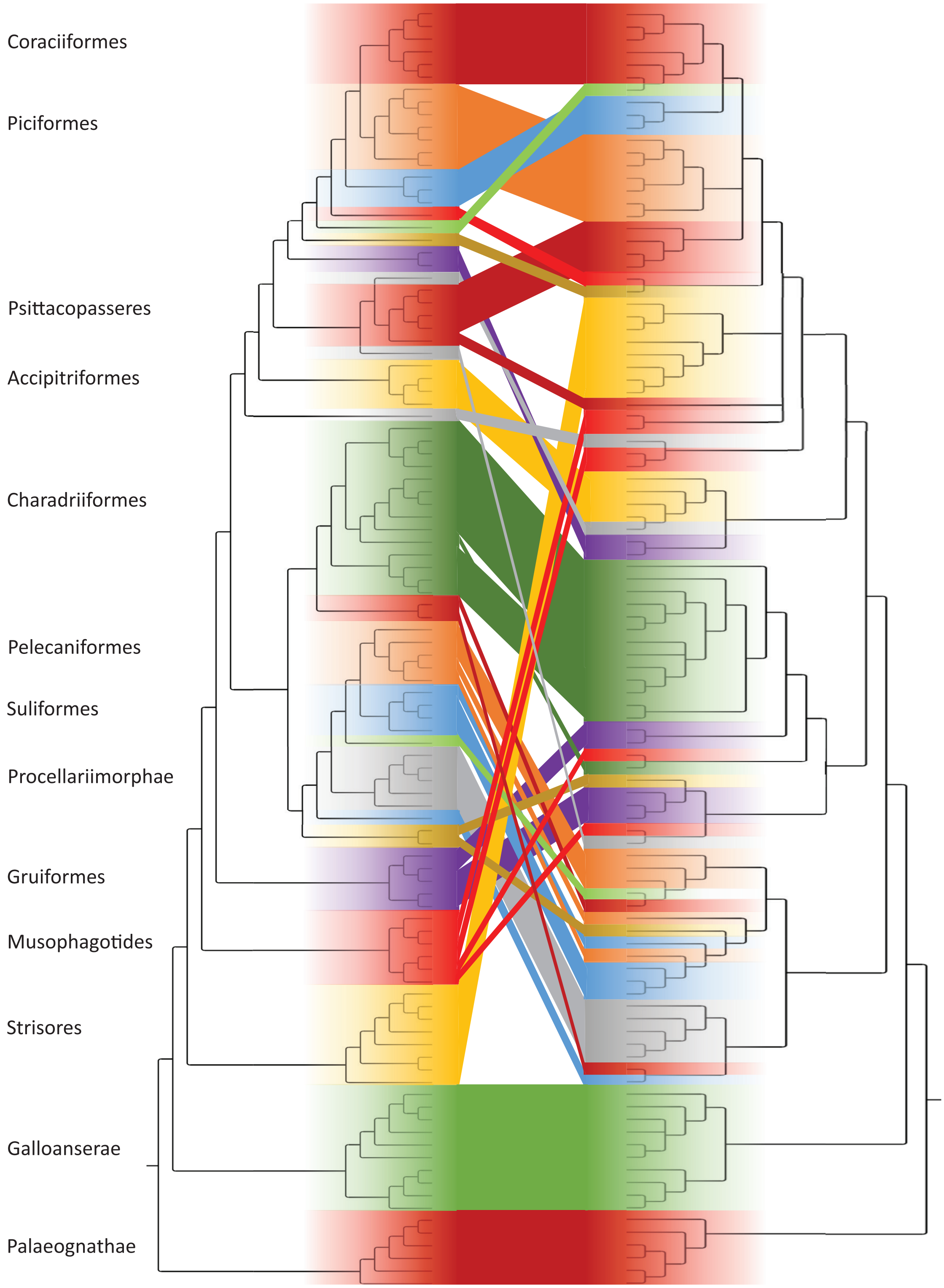

Figure 1. Continued

the 102 taxa in the 'Prum tree' and the analyses are based on these. The 1860 parsimony-informative characters were then grouped into different datasets according to category and anatomical region (NB: all numbering of characters in the text follows Livezey and Zusi 2006). In the first dataset (A) we divided the characters into osteological (nos. 1–2451), myological (nos. 2452–2708), and miscellaneous characters (nos. 2709–2954), respectively. We then defined two other datasets using only the osteological characters, which were divided according to anatomical regions. In dataset B, we contrasted cranial osteological characters (nos. 6–764) to postcranial ones (nos. 1–5 and 765–2451), and in dataset C we divided the osteological characters further into those belonging to the cranium (nos. 6–764), body (nos. 765–1344 and 1757–1964), wing (nos. 1345–1756), and leg (nos. 1965–2451), respectively.

Mesquite v.3.70 (Maddison and Maddison 2021) was used for graphical representation of characters mapped to the molecular 'Prum tree'.

We calculated the CI, RI, and RC for each character category and each data partition. In these statistics, we excluded all autapomorphic characters. We tested if observed differences in the fit of the various partitions differed significantly using the Mann–Whitney U test, which tests the null hypothesis of a 50% probability that the statistics of a randomly drawn member of the first population will exceed that of a member of the second population.

Evaluating the effect of combining morphological and molecular data

To study the effect of combining morphological and molecular characters in the phylogenetic analysis, we first extracted morphological characters that showed a low degree of homoplasy (CI = 0.5 or larger) across the 'Prum tree'. We then obtained morphological data for those 95 taxa that occur in both the dataset of Prum *et al.* (2015) and in another molecular-based, phylogenetic study (Hackett *et al.* 2008). We used this second, independent molecular dataset when evaluating the fit of the characters to avoid circularity in reasoning, which would have been the case if using the Prum *et al.* (2015) dataset. We added these 'low-homoplasy' morphological characters to the Hackett *et al.* (2008) molecular data and analysed the combined dataset with parsimony. The tree statistics for the tree obtained in the analysis of the combined dataset of morphological characters and the Hackett *et al.* (2008) molecular characters were then compared with those for a tree based on only molecular data.

Analysing the influence of doubtful homology assessments

In phylogenetic reconstructions using parsimony, ordered multistate characters may influence the phylogeny disproportionally if inferred step-matrices are based on incorrect homology assumptions. We used the myological dataset to address the potential influence of doubtful homology assessments by comparing the phylogenetic results (tree statistics and topologies) using the original coding of Livezey and Zusi (2006), with those after transforming all multistate characters to binary characters.

Identifying apomorphic characters for selected nodes

Finally, we also searched for characters that are apomorphic for certain higher-level groups of taxa in the 'Prum tree'. This was done by first mapping the morphological dataset to the 'Prum tree' and then filtering out all characters that have apomorphic states at the relevant nodes. We used CI = 0.5 or larger as a cut-off for the purpose to identify the least homoplastic character states.

RESULTS

Morphological homoplasy varies among categories of characters and anatomical regions

We investigated what parts of the avian morphology were most prone to convergent evolution by calculating tree statistics for both general categories of characters (osteological, myological, integumental, etc.) and several subgroupings of these categories. When mapped to the 'Prum tree', we found a generally high degree of homoplasy in all three major categories of characters with an overall CI of 0.38, RI of 0.50, and RC of 0.33 based on all 1860 parsimony-informative characters (Supporting information, Table S1). Of the 1860 variable characters, 985 (53%) showed a worse fit to the molecular tree than to the morphology tree, whereas 620 (33%) showed a better fit (Supplementary Material 1). After dividing the characters into the three major categories (dataset A), we found the myological characters to fit slightly worse (CI = 0.32, RI = 0.47, RC = 0.18) than the osteological characters (CI = 0.38, RI = 0.50, RC = 0.24), but only CI differed significantly ($P < .01$, Fig. 2; Supporting information, Table S2). Both the myological and the osteological categories fit was significantly worse than the miscellaneous characters (CI = 0.49, RI = 0.59, RC = 0.36) ($P < .001$, Fig. 2; Supporting information, Table S2).

Among the osteological characters (datasets B and C), the cranial characters fitted the 'Prum tree' significantly better (CI = 0.45, RI = 0.55, RC = 0.32) than the postcranial ones (CI = 0.36, RI = 0.48, RC = 0.21) ($P < .001$, Fig. 2; Supporting information, Table S2). The osteological body characters fitted the 'Prum tree' significantly worse (CI = 0.33, RI = 0.44, RC = 0.18) than those from the cranium (CI = 0.45, RI = 0.55, RC = 0.32), wing (CI = 0.39, RI = 0.49, RC = 0.23), and leg (CI = 0.38, RI = 0.52, RC = 0.24), respectively (Fig. 2; Supporting information, Table S2).

A large proportion (38%) of the 1860 morphological characters from Livezey and Zusi (2006) included herein were multistate, and about half of these (367) were analysed as ordered, i.e. a stepwise transformation series has been defined for them. In phylogenetic reconstructions using parsimony, treating characters as ordered means that the cost for a change from, e.g. 0 to 2 becomes two steps, while it only would be a single step if the character is treated as unordered. If the homology assessment is not correct, ordered multistate characters may be given too large a weight and thus influence the phylogenetic reconstruction. To address this concern, we analysed the 233 characters in the myological dataset, of which 85 were coded by Livezey and Zusi (2006) as multistate (52 ordered and 33 unordered). We transformed the 52 multistate ordered characters to binary characters and added them to the 181 other characters. By this the total number of myological characters increased to 817, but the results showed that including characters as ordered multistate or as binary did not significantly influence the overall tree statistics (Supporting information, Table S3), although the topologies differed (Supporting information, Fig. S2).

To test the robustness of our inference, we also compared our results for the 'Prum tree' with those based on another molecular phylogeny of birds (Stiller *et al.* 2024). After pruning the Stiller *et al.* (2024) and Prum *et al.* (2015) phylogenetic trees to include the same 98 taxa, the topologies of the two trees were fairly similar and both differed considerably from the topology of the 'Livezey & Zusi tree'. We found that the differences in summary statistics for the two molecular trees were negligible (Supporting information, Table S4), suggesting that the results of our comparisons with the morphological tree would be similar also if we used another phylogeny derived from a large amount of molecular data.

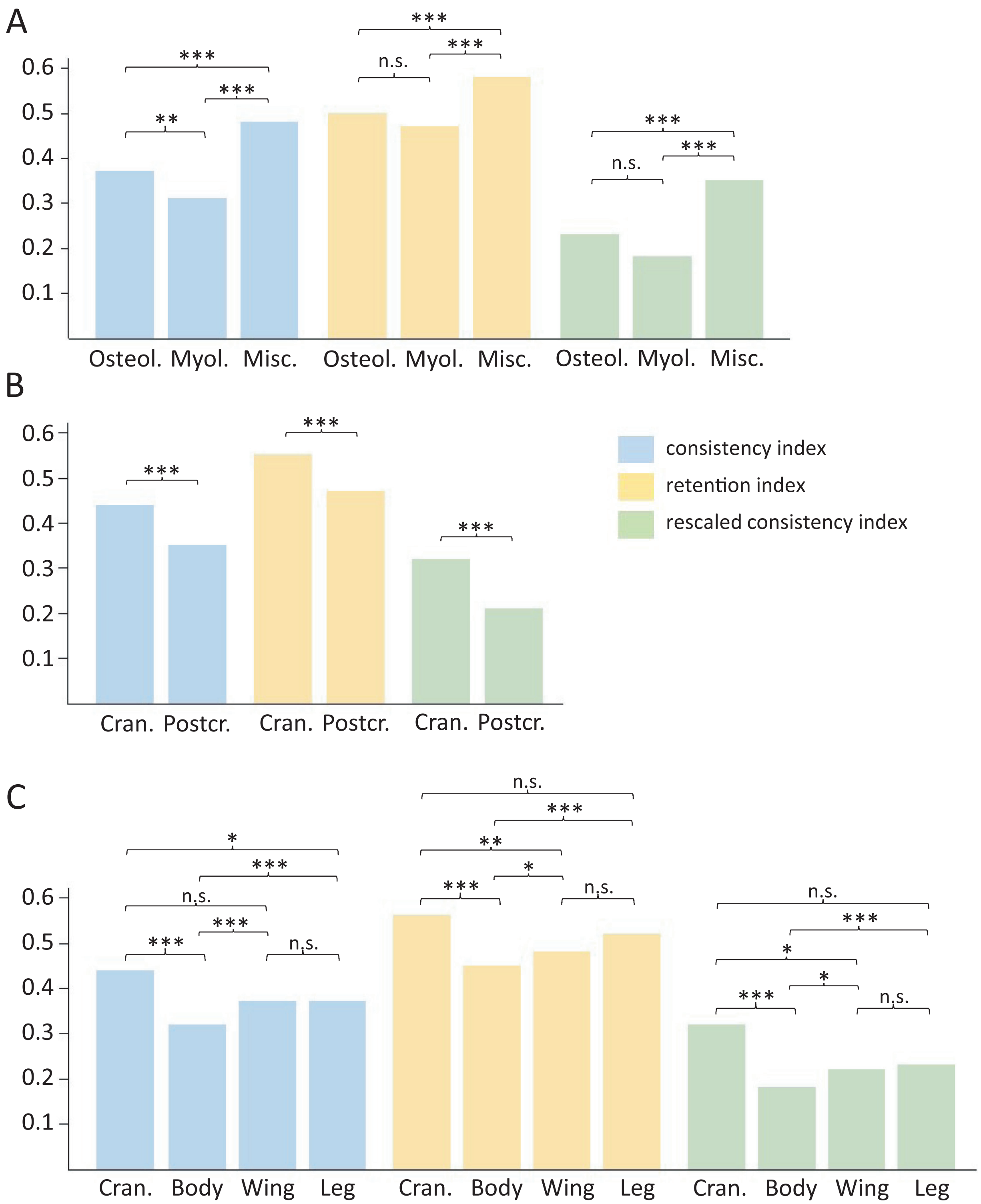

Figure 2. Results of the Mann–Whitney U test of differences in the observed mean values for the consistency, retention, and rescaled consistency indices. Three different groupings of data partitioning are shown: (A) comparison between the main categories of morphological characters, (B) comparison of cranial and postcranial osteological characters, and (C) comparison of osteological characters from different anatomical regions. See Supporting information, Tables S1 and S2 for detailed statistics. *** $P < .001$, ** $P < .01$, * $P < .05$, n.s. no statistically significant difference.

The effect of adding morphological characters to a molecular dataset

To evaluate if the addition of morphological information to an analysis of a molecular dataset improves the tree statistics we pruned the dataset of Hackett *et al.* (2008) to include the same 95 taxa for which we had morphological data. This molecular dataset, which consists of 55 330 bp, is completely different from the dataset of Prum *et al.* (2015), and was chosen to avoid

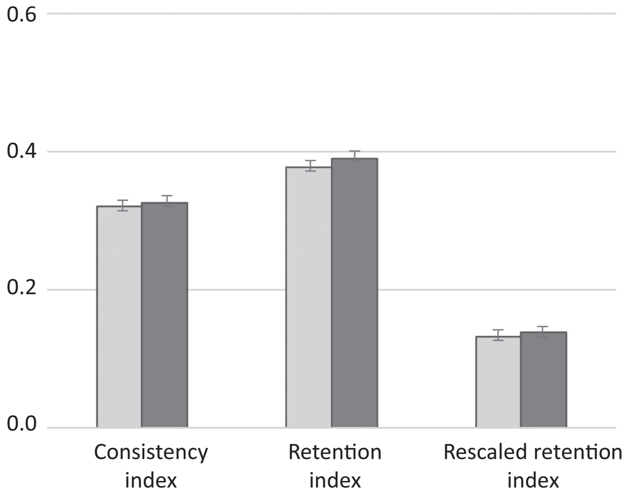

Figure 3. Tree statistics for the pruned (see Material and methods) molecular dataset of Hackett *et al.* (2008) before (light grey) and after (dark grey) adding 664 'low-homoplasy' morphological characters of the Livezey and Zusi (2006) dataset. Whiskers show the standard error around the mean. The observed differences are not statistically different.

Table 1. Morphological support (measured as number of apomorphic characters) for some major clades of birds identified in the molecular analysis (cf. Fig. 1). The total dataset includes the 1860 characters that are variable among the 102 taxa in the analysis, and only characters with a CI = 0.5 or larger are included in the summary counts. Clade names refer to Prum *et al.* (2015) when available. Asterisks mark clades that are not recovered in the original morphology tree of Livezey and Zusi (2007). The number of characters that are unambiguous (= synapomorphic) is given (second column), along with the number of osteological characters (third column). The latter are those that have the best chance to be preserved in the fossil record.

	Number of apomorphic characters		
	Total	**Unambiguous**	**Osteological**
Coraciiformes	2	0	1
Piciformes	11	3	10
Psittacopasseres *	0	-	-
Australaves *	0	-	-
Accipitriformes *	8	0	6
Charadriiformes *	3	0	2
Pelecaniformes *	1	0	1
Suliformes *	12	2	9
Procellariimorphae	11	8	6
Gruiformes *	0	-	-
Musophagotides *	0	-	-
Strisores	10	1	7
Galloanserae	15	12	12

circularity in reasoning. We first analysed this dataset alone after pruning it to include the 95 taxa for which we also have morphological data. We then re-ran the analysis after having added those 664 morphological characters that map to the 'Prum tree' with a low degree of homoplasy (i.e. a consistency index of 0.5 or higher). When comparing the statistics for the molecular and combined trees we did not find their fit significantly better to those of the 'Prum tree' (Fig. 3). We also compared the tree statistics for the combined dataset when mapped to the original tree of Hackett *et al.* (2008), pruned to include 95 taxa, and a new tree estimated from the combined dataset. We found that the tree fit statistics were almost identical for all these trees, albeit the tree lengths differed (Supporting information, Table S5).

Identifying apomorphic morphological characters that define basal clades of birds

To find morphological characters that may prove to be apomorphic for certain higher-level groups we mapped all them to the 'Prum tree' and then filtered out those that were the least homoplastic (having a CI of 0.5 or larger). However, we found that only a few of the selected nodes have any apomorphic characters, and several of them lacked such a character entirely (Table 1; Supporting information, Table S6). Moreover, those nodes at which apomorphic characters indeed occurred also do not have many characters (on average eight characters).

DISCUSSION

The observed worse fit of the osteological characters vs. those from miscellaneous 'soft-tissues' (mostly integument, feathers, intestinal, respiratory, syrinx, heart karyotype, urogenital, nervous, and reproductive organs) contradicts the results of Sansom and Wills (2017) who found a significantly better fit of 'hard' characters than 'soft' characters when mapped to an independent molecular tree.

As described, the 1860 parsimony-informative characters of Livezey and Zusi (2006) show a considerable degree of homoplasy (CI = 0.38, RI = 0.50, RC = 0.24) when fitted to the 'Prum tree'. The observation that the tree statistics are worse for the molecular tree than for the 'Livezey & Zusi tree' may indicate either that the molecular tree does not show the correct phylogenetic relationships, or that there is problem with the homology assumptions of many characters. However, we do consider the tree of Prum *et al.* (2015) to represent the evolutionary history of birds reasonably well as it is in general agreement with the phylogenetic relationships obtained in several other studies of molecular data (Ericson *et al.* 2006, Hackett *et al.* 2008, McCormack *et al.* 2013, Jarvis *et al.* 2014). This suggests that the major problem lies in the homology assumptions for many characters and character states across taxa, which is further supported by the observation that these characters show only marginally better fit to the tree that was calculated from themselves (Livezey and Zusi 2007) than to the molecular tree (Supporting information, Fig. S3, Table S7). However, when applying the homology assumptions and taxonomic distributions as defined by Livezey and Zusi (2006) these characters still contain a phylogenetic signal that is stronger than the noise in the dataset. This is evident from comparing with the tree statistics obtained when mapping them to 1000 random trees (constrained only to maintain paleognathous, neognathous, and neoavian taxa as monophyletic). For these random trees, CI ranges between 0.11 and 0.12, RI between 0.20 and 0.26, and RC between 0.02 and 0.03 (Supporting information, Table S8), which is considerably lower than when mapped to both the morphological and molecular trees.

How good are the homology assessments?

The obviously high degree of homoplasy in the morphological dataset suggests that these characters (as defined by Livezey and Zusi 2006), provide rather limited information about higher-level relationships in birds. However, one may then look at these results from the opposite side, i.e. how can the homology assumptions best be formulated to fit the phylogeny of birds? This methodological approach is not feasible to apply for all 1860 characters and here we can only discuss a few examples. Supplementary Material 1 lists the fit for all individual characters when mapped to both the morphological 'Livezey & Zusi tree' and the molecular 'Prum tree'. Of the 403 characters that showed perfect fit to (CI = 1.0) for the morphological tree, as many as 173 (43%) were found to be homoplastic when mapped to the molecular tree (CI ranging between 0.25 and 0.86). This reflects the marked topological differences between the two trees. However, it also indicates the difficulty of assessing the homology of character states and many of the characters showing the largest differences in fit between the trees are related to convergently evolved adaptations. For example, the foot-propelled divers (loons and grebes) are sisters in the morphological tree but distantly related in the molecular tree, the same goes for the flightless ratites. Loons and grebes share nine unambiguous synapomorphies according to the homology assessments of Livezey and Zusi (2006). Six of these concern osteological adaptations in the leg, all of which appear homoplastic when mapped to the 'Prum tree'. The ratites, in turn, are monophyletic in the morphological tree, leaving the likewise paleognathous tinamous outside. In the molecular tree, the tinamous are instead nested within the ratites and the 53 unambiguous morphological characters supporting ratite monophyly all become homoplastic when mapped to the 'Prum tree' (Supporting information, Fig. S4). About half of these are directly related to the loss of flight capacity in the ratites, which from the molecular tree is inferred to have happened at least twice (assuming that the origin of flight in tinamous did not evolve independently to that in all other birds). Many of the other worst fits of morphological characters to the molecular tree involve the ciconiiform, pelecaniform, and procellariiform birds, for which molecular analyses consistently have suggested phylogenetic relationships that are radically different from those traditionally inferred from morphology. The same is true for relationships among the higher landbirds, most notably in the position of the passerines and the non-monophyly of diurnal raptors and allies.

Many topological differences between the morphology-based 'Livezey & Zusi tree' and the molecular-based 'Prum tree' (Fig. 1B) involve groups assumed to have evolved similar lifestyles independently. The anatomical constraints that follow from birds' adaptations to flight have considerably decreased their amount of morphological variation, which contributes to the fact that selection for a certain lifestyle often results in superficially similar morphologies. Consequently, many similarly adapted groups, such as, e.g., foot-propelled divers (grebes and loons), wing-propelled divers (auks and penguins), and diurnal raptors (accipitrids and falconids), have independently developed morphologies that have led avian systematists to assume they are closely related, while recent molecular analyses have found them to belong to different clades that diverged more than 60 Mya (cf. Prum *et al.* 2015: fig. 1).

Combining morphological and molecular characters does not improve tree statistics

It has been shown that adding morphological information to a phylogenetic analysis of molecular data may increase the phylogenetic signal (e.g., Lee 2004, Lee and Camens 2009). However, we did not observe such an improvement of the phylogenetic signal after adding the 'best' morphological characters to the molecular dataset of Hackett *et al.* (2008). Applying a strict cut-off with a CI higher than 0.5, as used herein, could possibly improve the tree statistics, but such a low-homoplasious morphological dataset seems unlikely to be found in reality.

Few apomorphic characters define basal clades of birds

The paleontological record of modern birds is rich and many fossils have been possible to assign to family-level (either as members of the crown-group or of the stem-group). However, at taxonomic levels above the family the problem becomes much larger. Many fossil birds have thus been placed in extinct families of unclear taxonomic affinities (cf. Olson 1985). At least in theory, these fossils should be able to refer to higher-level categories if we knew what morphology to expect from the ancestors of the modern groups (Mayr 2017). In theory we could obtain information about this by reconstructing ancestral states for the morphological characters of Livezey and Zusi (2006) at selected nodes in the 'Prum tree'. However, when doing this for basal nodes that are strongly supported by molecular data (Fig. 1; Table 1) we did not find any apomorphic character with a CI of 0.5 or larger for several of these (Table 1; Supporting information, Table S8). Our analysis thus shows that the data of Livezey and Zusi (2006) includes few morphological characters that unambiguously support groups of families or orders. The phylogenetic information from avian morphology [provided that the Livezey and Zusi (2006) dataset is a good representation of the available variation] thus seems to be of rather limited use in reconstructing higher-level relationships, which raises questions about our ability to unambiguously identify a fossil bird as the direct ancestor of two or more extant families. For example, molecular data suggest parrots and passerines to be sister groups (forming the clade Psittacopasseres), which is surprising as they differ morphologically and have never been suggested as sisters before. None of the morphological characters of Livezey and Zusi (2006) are apomorphic (with a CI = 0.5 or larger) for the Psittacopasseres. It thus seems highly unlikely one would be able to recognize the ancestor of this clade in the fossil record.

Convergent evolution and large-scale biogeographic patterns

As noted above, among the more noteworthy topological differences between the morphological 'Livezey & Zusi tree' and the molecular 'Prum tree' are the relative positions of certain groups with similar adaptations, e.g. grebes and loons (foot-propelled divers), auks and penguins (wing-propelled divers), and accipitrids and falconids (diurnal raptors). This observation bears on two important issues in avian evolution. First, it testifies to the problem of ascertain homologies of morphological characters related to feeding and locomotory adaptations in birds, in agreement with numerous other examples where similar morphologies have evolved independently in two or more lineages (e.g. McCracken *et al.* 1999, Fain and Houde 2004, Irestedt *et al.* 2004, Ericson 2012). It seems symptomatic that the 178 characters in the 'Miscellaneous' category show an average better best fit to the molecular phylogeny than characters in the other categories (Supporting information, Table S1). The 'Miscellaneous' category includes characters related to, e.g. integument, feathers, intestinal, respiratory, syrinx, heart karyotype, urogenital, nervous, and reproductive organs (Livezey and Zusi 2006), which are not as involved in feeding and locomotory adaptations as are many osteological and myological characters.

The second issue about the evolution of birds highlighted by the results of this study is the question of where in the world the earliest radiations within Neoaves took place. Several neoavian groups that have developed similar morphological adaptations in parallel belong to clades that split very long ago, some perhaps even around the K/Pg border at 66 Mya (cf. Wu *et al.* 2024). The very old age of these clades, along with the many examples of parallel evolution, suggest parallelisms in avian guild evolution in different parts of the world, resulting in the independent morphological adaptations to similar ecological niches. Such a scenario has previously been suggested for the two largest terrestrial radiations in birds, Afroaves and Australavis (Ericson 2012), but may also involve birds adapted to aquatic life-styles. Despite their potentially high mobility, the distribution of modern birds shows considerable biogeographic structure at all taxonomic levels, from the ordinal level to the intraspecific. This is true not only for terrestrial groups, but also for many aquatic and even pelagic groups, regardless of whether the species are predominantly stationary, short-distance migrators, or long-distance migrators. Today's biogeographic structure is explained partly by the local extinction of some groups (e.g. Mourer-Chauviré 1982, Mayr 2017), but there is no doubt that general patterns of geographic distribution in birds reflect their earliest evolution (e.g. Cracraft 2001, Barker *et al.* 2002, Ericson *et al.* 2002). However, a detailed analysis of the geographical origins of the hypothesized ecological guilds is beyond the scope of the present study.

The morphological characters defined by Livezey and Zusi (2006) carry important phylogenetic information, despite this convergent morphological evolution following the adaptation to similar life-styles in different lineages of birds has made it difficult to ascertain the homologies. We believe our study is the beginning of a large-scale re-evaluation of the morphological variation in birds in the light of phylogenetic information from independent data. Such work would constitute a perfect example of 'reciprocal illumination' [as defined by Hennig (1950, 1966)], that surely will spark a renewed interest in studies of morphological adaptation at higher taxonomic levels in birds, as well as proving useful in paleontological studies of fossils belonging to taxa distantly related to modern avian families.

SUPPLEMENTARY DATA

Supplementary data is available at *Biological Journal of the Linnean Society* online.

ACKNOWLEDGEMENTS

We thank Xuexing Zheng for performing some of the statistical tests and for discussing the results of the study. We also acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and the SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by the National Natural Science Foundation of China (grants no. NSFC32020103005 and U23A20162 to Y.Q.) and the Swedish Research Council (grant no. 621-2017-3693 to P.E.).

DATA AVAILABILITY

The data analysed in this study is available in Dryad (doi: 10.5061/dryad.r2280gbmz).

REFERENCES

Archie JW. Homoplasy excess statistics and retention indices: a reply to Farris. *Systematic Zoology* 1990;**39**:169–74. https://doi.org/10.2307/2992454

Baker RH, Yu X, De Salle R. Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. *Molecular Phylogenetics and Evolution* 1998;**9**:427–36.

Barker FK, Barrowclough GF, Groth JG. A phylogenetic hypothesis for passerine birds: taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data. *Proceedings Biological Sciences* 2002;**269**:295–308. https://doi.org/10.1098/rspb.2001.1883

Cracraft J. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. *Proceedings Biological Sciences* 2001;**268**:459–69. https://doi.org/10.1098/rspb.2000.1368

Ericson PGP. Evolution of terrestrial birds in three continents: biogeography and parallel radiations. *Journal of Biogeography* 2012;**39**:813–24. https://doi.org/10.1111/j.1365-2699.2011.02650.x

Ericson PGP, Christidis L, Cooper A *et al.* A Gondwanan origin of passerine birds supported by DNA sequences of the endemic New Zealand wrens. *Proceedings Biological Sciences* 2002;**269**:235–41. https://doi.org/10.1098/rspb.2001.1877

Ericson PGP, Anderson CL, Britton T *et al.* Diversification of Neoaves: integration of molecular sequence data and fossils. *Biology Letters* 2006;**2**:543–7. https://doi.org/10.1098/rsbl.2006.0523

Fain MG, Houde P. Parallel radiations in the primary clades of birds. *Evolution* 2004;**58**:2558–73. https://doi.org/10.1111/j.0014-3820.2004.tb00884.x

Farris JS. Methods for computing Wagner trees. *Systematic Zoology* 1970;**19**:83–92. https://doi.org/10.2307/2412028

Farris JS. The retention index and the rescaled consistency index. *Cladistics* 1989;**5**:417–9. https://doi.org/10.1111/j.1096-0031.1989.tb00573.x

Fitch WM. Toward defining the course of evolution: minimum change for a specified tree topology. *Systematic Zoology* 1971;**20**:406–16. https://doi.org/10.2307/2412116

Gatesy J, Arctander P. Hidden morphological support for the phylogenetic placement of *Pseudoryx nghetinhensis* with bovine bovids: a combined analysis of gross anatomical evidence and DNA sequences from five genes. *Systematic Biology* 2000;**49**:515–38. https://doi.org/10.1080/10635159950127376

Gatesy J, Amato G, Norell M *et al.* Combined support for wholesale taxic atavism in gavialine crocodylians. *Systematic Biology* 2003;**52**:403–22. https://doi.org/10.1080/10635150390197037

Goloboff PA. *Refining Phylogenetic Analyses: Phylogenetic Analysis of Morphological Data: Volume 2 - Species and Systematics.* Boca Raton: Taylor & Francis Ltd., 2022.

Hackett SJ, Kimball RT, Reddy S *et al.* A phylogenomic study of birds reveals their evolutionary history. *Science* 2008;**320**:1763–8. https://doi.org/10.1126/science.1157704

Hennig W. *Grundzüge einer Theorie der phylogenetischen Systematik.* Berlin: Deutscher Zentralverlag, 1950.

Hennig W. *Phylogenetic Systematics.* Urbana, IL: University of Illinois Press, 1966.

Hull DL. *Science as a Process.* Chicago, IL: University of Chicago Press, 1988.

Irestedt M, Fjeldså J, Ericson PGP. Phylogenetic relationships of woodcreepers (Aves: Dendrocolaptinae) - incongruence between molecular and morphological data. *Journal of Avian Biology* 2004;**35**:280–8. https://doi.org/10.1111/j.0908-8857.2004.03234.x

Jarvis ED, Mirarab S, Aberer AJ *et al.* Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 2014;**346**:1320–31. https://doi.org/10.1126/science.1253451

Kluge AG. A concern for evidence, and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology* 1989;**38**:7–25. https://doi.org/10.2307/2992432

Kluge AG, Farris JS. Quantitative phyletics and the evolution of anurans. *Systematic Zoology* 1969;**18**:1–32. https://doi.org/10.2307/2412407

Lee MSY. Molecular and morphological datasets have similar numbers of relevant phylogenetic characters. *Taxon* 2004;**53**:1019–22.

Lee MSY, Camens AB. Strong morphological support for the molecular evolutionary tree of placental mammals. *Journal of Evolutionary Biology* 2009;**22**:2243–57. https://doi.org/10.1111/j.1420-9101.2009.01843.x

Livezey BC, Zusi RL. Higher-order phylogenetics of modern Aves based on comparative anatomy. *Netherlands Journal of Zoology* 2001;**51**:179–205. https://doi.org/10.1163/156854201x00260

Livezey BC, Zusi RL. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. I. Methods and characters. *Bulletin of the Carnegie Museum of Natural History* 2006;**37**:1–556.

Livezey BC, Zusi RL. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. *Zoological Journal of the Linnean Society* 2007;**149**:1–95. https://doi.org/10.1111/j.1096-3642.2006.00293.x

Maddison WP, Maddison DR. *Mesquite: a Modular System for Evolutionary Analysis. Version 3.70.* http://www.mesquiteproject.org, 2021 (22 April 2022, date last accessed).

Magallón SA. Dating lineages: molecular and paleontological approaches to the temporal framework of clades. *International Journal of Plant Sciences* 2004;**165**:S7–S21. https://doi.org/10.1086/383336

Mayr G. Avian higher-level phylogeny: well-supported clades and what we can learn from a phylogenetic analysis of 2954 morphological characters. *Journal of Zoological Systematics and Evolutionary Research* 2008;**46**:63–72. https://doi.org/10.1111/j.1439-0469.2007.00433.x

Mayr G. *Avian Evolution: the Fossil Record of Birds and its Paleobiological Significance.* Chichester, UK: Wiley, 2017.

McCormack JE, Harvey MG, Faircloth BC *et al.* A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. *PLoS One* 2013;**8**:e54848. https://doi.org/10.1371/journal.pone.0054848

McCracken KG, Harshman J, McClellan DA *et al.* Data set incongruence and correlated character evolution: an example of functional convergence in the hind-limbs of stifftail diving ducks. *Systematic Biology* 1999;**48**:683–714. https://doi.org/10.1080/106351599259979

Mickevich MF, Lipscomb D. Parsimony and the choice between different transformations for the same character set. *Cladistics* 1991;**7**:111–39. https://doi.org/10.1111/j.1096-0031.1991.tb00028.x

Mourer-Chauviré C. Les oiseaux fossiles des phosphorites du quercy (éocène supérieur a oligocène supérieur): implications paléobiogéographiques. *Geobios* 1982;**15**:413–26. https://doi.org/10.1016/s0016-6995(82)80130-7

Murrell A, Campell NJH, Barker SC. A total-evidence phylogeny of ticks provides insights into the evolution of life cycles and biogeography. *Molecular Phylogenetics and Evolution* 2001;**21**:244–58.

Nixon K, Carpenter J. On simultaneous analysis. *Cladistics* 1996;**12**:221–41.

Olson SLA. A critique of Cracraft's classification of birds. *The Auk* 1982;**99**:733–9.

Olson SL. The fossil record of birds. In: Farner DS, King JR, Parkes KC, (eds), *Avian Biology*, **8**. New York: Academic Press, 1985, 79–238.

Parham JF, Donoghue PCJ, Bell CJ *et al.* Best practices for justifying fossil calibrations. *Systematic Biology* 2012;**61**:346–59. https://doi.org/10.1093/sysbio/syr107

Patterson C. Significance of fossils in determining evolutionary relationships. *Annual Review of Ecology and Systematics* 1981;**12**:195–223. https://doi.org/10.1146/annurev.es.12.110181.001211

Prum RO, Berv JS, Dornburg A *et al.* A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 2015;**526**:569–73. https://doi.org/10.1038/nature15697

Rieppel O. The philosophy of total evidence and its relevance for phylogenetic inference. *Papéis Avulsos de Zoologia* 2005;**45**:8.

Sangster G, Braun EL, Johansson US *et al.* Phylogenetic definitions for 25 higher-level clade names of birds. *Avian Research* 2022;**13**:100027. https://doi.org/10.1016/j.avrs.2022.100027

Sansom RS, Wills MA. Differences between hard and soft phylogenetic data. *Proceedings Biological Sciences* 2017;**284**:20172150. https://doi.org/10.1098/rspb.2017.2150

Stiller J, Feng S, Chowdhury AA *et al.* Complexity of avian evolution revealed by family-level genomes. *Nature* 2024;**629**:851–60. https://doi.org/10.1038/s41586-024-07323-1

Swofford DL. *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods).* Sunderland, MA: Sinauer Associates, Inc., 2000.

Wortley AH, Scotland RW. The effect of combining molecular and morphological data in published phylogenetic analyses. *Systematic Biology* 2006;**55**:677–85. https://doi.org/10.1080/10635150600899798

Wu S, Rheindt FE, Zhang J *et al.* Genomes, fossils, and the concurrent rise of modern birds and flowering plants in the Late Cretaceous. *Proceedings of the National Academy of Sciences USA* 2024;**121**:2319696121.