Early diversification of avian limb morphology and the role of modularity in the locomotor evolution of crown birds

Abstract

High disparity among avian forelimb and hind limb segments in crown birds relative to non-avialan theropod dinosaurs, potentially driven by the origin of separate forelimb and hind limb locomotor modules, has been linked to the evolution of diverse avian locomotor behaviors. However, this hypothesized relationship has not been quantitatively investigated in a phylogenetic framework. We assessed the relationship between the evolution of limb morphology and locomotor behavior by comparing a numerical proxy for locomotor disparity to morphospace sizes derived from a dataset of 1241 extant species. We then estimated how limb disparity accumulated during the crown avian radiation. Lastly, we tested whether limb segments evolved independently between each limb module using phylogenetically informed regressions. Limb disparity increased significantly with locomotor disparity after accounting for clade age and species richness. We found that forelimb disparity accumulated rapidly early in avian evolution, whereas hind limb disparity accumulated later, in more recent divergences. We recovered little support for strong correlations between forelimb and hind limb morphology. We posit that these findings support independent evolution of locomotor modules that enabled the striking morphological and behavioral disparity of extant birds.

Key Words: Locomotor modules; morphological disparity; avian macroevolution

Introduction

Extant birds (Aves) are among the most speciose and ecologically diverse clades of tetrapods. This diversity is reflected in the array of locomotor behaviors birds employ: birds use their wings and legs to fly, walk, run, climb, swim, and dive. Behavioral variation has been linked to a shift in limb function and neural control early in the evolution of birds that resulted in the simultaneous origin of three separate functional units, termed locomotor modules (Gatesy and Dial 1996). In the evolution of these modules, the forelimb was first co-opted for flight, the hind limb and tail were subsequently decoupled from their shared role in terrestrial locomotion, and then the tail was integrated into the flight apparatus (Gatesy and Dial 1996). The evolution of these three modules in Avialae is well documented in the fossil record by musculoskeletal evolution in the avialan dinosaur hind limb and tail and shifts in hind limb scaling from non-avialan theropod dinosaurs to crown birds (Gatesy 1991, Hutchinson 2001a; 2001b; Gatesy 2002; Hutchinson 2002; Clarke et al. 2006; Clarke and Middleton 2008; Hutchinson and Allen 2009; Allen et al. 2013; Dececchi and Larsson 2013). Locomotor modularity was hypothesized to have allowed birds to evolve novel locomotor strategies by differentially elaborating individual modules. Specifically, the morphological and functional evolutionary trajectory of each module can be modified in different permutations, enabling a wider array of locomotor behaviors in birds relative to their non-avialan theropod ancestors (Gatesy and Dial 1996).

Forelimb and hind limb proportions are far more variable in crown birds than in non-avialan theropod dinosaurs (Gatesy and Middleton 1997; Carrano and Sidor 1999; Middleton and Gatesy 2000; Benson & Choiniere 2013; Mitchell & Mackovicky 2014). One explanation for this pattern is a hypothesized link between limb morphology and ecological traits, including flight capabilities and limb use. However, establishing this relationship has proven challenging, in part, due to the many factors influencing limb morphology, from allometric and phylogenetic factors (Doube et al. 2012; Nudds et al. 2013; Stoessel et al. 2012; Kilbourne 2013; Wang and Clarke 2014) to biomechanical factors such as maintaining stability and low energetic costs during movement (Zeffer et al. 2003; Stoessel et al. 2012). Additionally, some locomotor behaviors influence limb morphology more strongly than others, in particular wading or different types of aquatic locomotion (Zeffer et al. 2003; Hinić-Frlog and Motani 2010; Wang and Clarke 2014; Baumgart et al. 2021). Thus, broad flight-style categories and aerodynamic variables (e.g., wing loading, aspect ratio) can fail to predict forelimb morphology and wing shape across large phylogenetic samples (e.g., Nudds et al. 2007; Taylor and Thomas 2014; Wang and Clarke 2014; 2015; Karoullas and Nudds 2021). That said, recent multivariate approaches have yielded more compelling results: a combination of multiple morphological characters across the body correctly predicted discretized categories of extant bird ecology (Mitchell and Makovicky 2014), and high-density morphometric data showed how limb shape covaries with ecology more strongly than limb size (Orkney et al. 2021). In addition to direct links between limb morphology and ecology, the locomotor modularity hypothesis (Gatesy and Dial 1996) predicts a positive correlation between variation in limb morphometrics and locomotor disparity that has not yet been evaluated.

Evolutionary dynamics of differential elaboration between locomotor modules is also poorly known. For example: How do limb measurements change when lineages evolve new ways of moving? Do limb modules covary due to selective pressure on whole-organism performance (see Clarke and Middleton 2008; Heers and Dial 2015; Ornkey et al. 2021)? Older comparisons of limb morphology and locomotor behavior in extant avian clades usually considered a single limb pair or analyzed each limb separately (e.g., Zeffer et al. 2003; Nudds et al. 2012; Stoessel et al. 2012; Wang and Clarke 2014). Scenarios that can produce covariation between locomotor modules, such as cooperative limb function perpetuated by mutually reinforcing selection or functional tradeoffs due to contrasting selection between individual limb functions, are of particular interest because they channel how locomotor modules can differentially elaborate. Multiple examples of cooperative use of forelimbs and hind limbs, specifically in juveniles that are non-volant, were previously discussed (see Dial 2003b; Heers et al. 2014; Dial et al. 2015). Cooperative function between limbs, particularly when negotiating 3D terrain, likely played a prominent role in shaping avian limb evolution during and after the origin of flight (Dial 2003b; Clarke and Middleton 2008; Dial et al. 2011; Heers et al. 2014).

Despite the potential importance of coevolution between locomotor modules in birds, until recently, there have been few quantitative comparisons between the forelimb and hind limb modules across all birds (but see Orkney et al. 2021). Rather, studies have considered morphological differences among different ecological guilds, the effect of flight loss on limb evolution within single clades (Livezey 1988; 1989; 2003), and comparisons among different locomotor behaviors within a lineage (e.g., Mitsuo 1955; Norberg 1979; Miles and Ricklefs 1984; Miles et al. 1987; Mitchell and Makovicky 2014; Hertel et al. 2015). Other work has showed that the ratio of forelimb to hind limb muscle mass negatively covaries with ecology between ground dwelling and aerial specialists (Heers and Dial 2015), inertial properties of the forelimb and hind limb scale differently with body mass and individual limb sizes (Kilbourne 2013), forelimbs and hind limbs cooperatively interact during takeoff and landing (Provini et al. 2012; Provini et al. 2014), and that island populations tend to possess decreased flight muscle mass and longer legs compared to mainland relatives (Wright et al. 2016). Taken together, these examples show that potential correlations between limbs should be considered when exploring macroevolutionary patterns in avian limb proportions, because selective pressures on the whole organism may affect differential elaboration between modules.

Here, we present an assessment of the evolution of avian forelimb and hind limb segments across a large dataset (N = 2549 measurements for 1241 species) of crown birds, with increased sampling in previously underrepresented subclades. Specifically, we test the following hypotheses about limb evolution in birds: i) limb disparity and locomotor behavior disparity are correlated at the major clade level and ii) covariation within hind limb and forelimb modules is stronger than that between modules. We first examine morphospace occupancy for each limb in each major subclade and investigate potential relationships among diversity in locomotor behavior and morphological evolution of limb segments. We predict that larger morphospace sizes will be positively correlated with locomotor disparity. We evaluate whether birds established their forelimb and hind limb morphospaces early in their evolutionary history. Furthermore, we test for potential evolutionary covariation within and among locomotor modules by examining evolutionary rates in limb element lengths and phylogenetically informed covariation of segment lengths between limbs. If modules are evolutionarily linked, we predict similar evolutionary covariation within limbs and between homologous elements (e.g., humerus and femur), whereas weaker covariation between limbs would support differential elaboration between avian locomotor modules.

Methods

Morphometric data set assembly. To investigate the evolution of avian limb morphospace, we obtained 2549 raw lengths (in mm) of forelimb and hind limb segments (i.e., humerus, radius/ulna, carpometacarpus, femur, tibiotarsus, and tarsometatarsus; see Fig. 1) for 1241 species of extant birds from all major avian subclades, as well as for 4 recently extinct taxa (see Supplementary Methods). Forelimb and hind limb data for 1121 of these taxa, as well as the 4 extinct species, were taken from previous studies (Gatesy and Middleton 1997; Middleton and Gatesy 2000; Hinić-Frlog and Motani 2010; Nudds et al. 2013; Mitchell and Makovicky 2014). Data for 120 new extant taxa were added through measurement of museum specimens to supplement underrepresented subclades (e.g., Coraciimorphae N = 30, Aequornithia N = 29, and Galloanseres N = 19 specimens; see Supplementary Methods). Raw measurements for each limb of multiple individuals of the same species were subsequently used in comparative analyses to account for measurement error, and species means were used to test hypotheses of limb evolution (see below). All statistical analyses were carried out in R v. 3.4 (R core team 2017). Graphics were created using the R package ggplot2 (Wickham 2009).

Phylogenetic tree construction. We used a supertree of avian taxa (Burleigh et al. 2015) pruned to the taxon set with measurement data to estimate the evolution of limb-proportion disparity-through-time, as well as to test evolutionary covariation between limbs. We employed a modified version of the Burleigh et al. (2015) tree that was scaled to absolute time divergence estimates in millions of years from a recent genome-based avian phylogeny (Jarvis et al. 2014; Reide et al. 2016). We pruned the supertree using the R packages ape v. 3.3 and geiger v. 2.06 to include only taxa that had data from both limb pairs, resulting in a 1140 taxon tree (Paradis et al. 2004; Harmon et al. 2008). We also used another time-calibrated molecular phylogeny (Prum et al. 2015) to test if topology and sampling affected our results, pruning it to a tree comprising 183 taxa. In cases where the original trees included species that were not sampled (6 taxa for Burleigh et al. 2015; 56 taxa for Prum et al. 2015), we replaced those taxa with sampled members of the same genus. For 6 species from the Prum et al. (2015) tree congeneric taxa were not sampled. In these cases we replaced these tips with the next most closely related taxon in our dataset according to current phylogenetic hypotheses (see Moyle 2004; Alström et al. 2013; Ramirez et al. 2013; Garcia-Ramirez et al. 2014; Dufort 2016; Xia et al. 2016).

Quantifying avian limb morphospace. To visualize avian hind limb and forelimb morphospace, we took the species averages of log-transformed limb segment lengths. Next, to account for allometric scaling, we computed phylogenetic residuals under a Brownian motion (BM) model of trait evolution for each log-transformed trait regressed on body size. We then calculated relative limb segment lengths as the residuals of a phylogenetic regression on log body mass using the phyl.resid function in phytools (Revell 2012). For visualization purposes, we conducted separate principal components analyses (PCA) on residual forelimb (3 traits) and hind limb segment lengths (3 traits) and visualized each morphospace using PC1 and PC2 (we retained these PCs because they accounted for >95% for each limb module; see Fig. 2). To quantify clade-specific morphospace sizes, we calculated the volume of an N-dimensional convex hull circumscribing species traits (i.e. residual segment lengths; see above) using the convhulln function in the R package geometry v. 0.4.5. Although we originally defined 18 clades, convex hulls require at least 3 points for estimating 3-D volumes (for the hind limb, forelimb) and 6 points for 6-D volumes (for the overall limb morphospace). Therefore, we obtained overall limb morphospace volumes for 15 clades and limb-specific volumes for 16 clades. These morphospace volumes were used in understanding clade-specific patterns of morphospace occupancy and in calculating lineage density (see below), but we did not compare volumes directly in comparative analyses. To determine a clade's uniqueness in morphospace, we determined the non-overlapping volume of each clade using the voloverlap function of pavo (Maia et al. 2013).

Quantifying locomotor disparity. To characterize clade-specific locomotor disparity, we coded each clade for the presence or absence of several different locomotor behaviors using life history data in the Handbook of the Birds of the World (del Hoyo et al. 1992-2016). In total, we identified and scored ten unique forelimb-related behaviors and eleven hind limb-related behaviors in each clade (see Supplementary Methods). We then summed the number of behaviors as discrete integer scores of locomotor disparity for each clade. Although birds often perform locomotor behaviors considered outside their typical repertoire, we only included behaviors commonly used by each clade, as determined by descriptions of their life history. Importantly, our approach follows Taylor and Thomas (2014) in coding taxa (i.e. clades) as a mosaic of multiple locomotor behaviors rather than assigning taxa to single locomotor categories (e.g., see Baumgart et al. 2021). Caveats of our mosaic approach include i) all locomotor behaviors are treated as functionally equivalent (e.g., a clade with hovering and sallying has the same locomotor score as a clade with wing-propelled diving and soaring) and ii) evolutionary variation at the species level may be missed (e.g., if a species loses a locomotor behavior specific to its clade). Yet, given the evolutionary scale of our question (i.e., we want to know whether locomotor disparity is associated with limb morphological disparity at the clade level) and the lack of life history data for several of the species in our dataset (del Hoyo et al. 1992-2016), we feel our approach is warranted.

Testing the locomotor disparity hypothesis. To test our hypothesis that larger morphospace sizes will be associated with greater locomotor disparity, we took a phylomorphospace approach (Silauskas et al. 2008). This is because expansions in morphospace can be caused either by elevated rates of evolution in some clades (i.e. those that utilize a greater number of locomotor modes) or differences in mode of evolution that expand morphospace (e.g., morphological innovations associate with different locomotor behaviors). These two scenarios can be teased apart analytically by analyzing rates of evolution and lineage density, respectively (Sidlauskas et al. 2008, Benson Choiniere 2013). Briefly, we reconstructed ancestral states for each set of trait residuals using fastAnc in phytools (Revell 2012) and calculated morphological disparity as the sum of Euclidean distances between all adjacent nodes and tips of the phylogeny (Sidlauskas 2008). We then divided these morphological branch lengths by phylogenetic branch lengths and averaged values for each clade to derive clade-wise estimates of evolutionary rate. This approach conveniently provides single rate estimates for the fore and hind limb, rather than an 3 x 3 rate matrix for each, as in a multivariate BM model. To calculate lineage densities, we divided the sum of morphological branch lengths within a clade by its morphospace volume (Sidlauskas 2008). Since limb disparity calculated as the sum of morphological branch lengths is potentially sensitive to the number of lineages in a clade (Sidlauskas 2008), we also calculated limb disparity as the mean pairwise Euclidean distances among lineages in morphospace (Foote 1993), a metric that is independent of sample size. We finally tested for relationships between response variables (morphological disparity, evolutionary rate, lineage density) and locomotor scores using phylogenetic generalized least squares (PGLS) multiple regressions implemented in the R package phylolm (Ho and Ané 2014).

While our phylomorphospace approach takes into account branch lengths, clade age and species richness could also influence our results. For example, if limb morphology is evolving by a Brownian motion process, then we would expect a positive relationship between limb disparity and clade age. Thus, we wanted to account for these covariates in our analyses. We used stepwise multiple regression to determine the best-fitting model based on AIC scores. For the PGLS analyses, we pruned the Kimball et al. (2019) time-calibrated tree by collapsing the 18 major clades into individual terminals (see ESM for results using the Prum et al. 2015 tree). We natural log-transformed response variables prior to regression and assessed normality of residuals with quantile-quantile plots in R. To assess the effect of potential outliers, we removed influential taxa and re-fit PGLS models in the influ_phylm function of sensiPhy (Paterno et al. 2018).

Understanding the tempo and mode of evolution of avian limb traits. We evaluated whether birds established their morphospace of limb dimensions early in their evolutionary history, compared to expectations under a diffusive, Brownian motion model. Following Cooney et al. (2017), we first estimated ancestral states of relative limb segment lengths using fastAnc (Revell 2012). We then divided the tree into 1-My time bins and, for each bin, calculated disparity as the mean Euclidean distance among lineages present at that time (Foote 1993). We calculated these disparities for each limb module separately (e.g., among humerus, radius/ulna, and carpometacarpus relative lengths for the forelimb module). To test whether disparities were higher or lower than expected by chance, we simulated multivariate Brownian motion evolution 100 times with mvSIM (Clavel et al. 2015) and re-calculated disparity-through-time curves for each simulated dataset. We then determined the 95% confidence interval with the quantile function in R. To compare forelimb and hind limb disparity trends, we determined differences in disparity for each time bin and tested whether this difference fell outside the 95% CI of the difference calculated from trait simulations (see Dryad for R code).

We further tested if evolution in one limb is correlated with evolution in the other by estimating evolutionary rates and covariances among our six residual segment lengths using the ratematrix R package (Caetano & Harmon 2018). Since these traits are all expected to covary strongly with body size, we also analyzed raw segment lengths with log body mass included (N = 7 traits). To statistically compare rates and covariances, we calculated pairwise differences between them and computed Bayesian P values as the proportion of the posterior sample overlapping zero, using the helper function pMCMC in MCMCglmm (Hadfield 2010). We adjusted P values for multiple tests with the false discovery rate (FDR) metric with the p.adjust function in R.

Results

Limb morphospace comparisons. Forelimb morphospace was twice as large as hind limb morphospace (0.176 log mm^3 versus 0.061 mm^3; Fig. 2). Regions of morphospace unique to particular clades comprised a smaller proportion of total hind limb variation than in the forelimb, as clades overlapped more in hind limb morphospace than in forelimb morphospace (Fig. 2). Most clades had larger hind limb than forelimb morphospace volumes (12/16, 75%), with a few notable exceptions: Strisores (nightjars, hummingbirds, and swifts; 0.015 versus 0.004 log mm^3) and Paleognathae (tinamous, "ratites"; 0.044 versus 0.0006 log mm^3; Fig. 2). Strisores, Aequornithes, and Paleognathae dominated unique forelimb morphospace (0.055 log mm^3; 97.3% of unique volume) due to a combination of shorter humeri (hummingbirds, swifts, treeswifts, and penguins), longer carpometacarpi (hummingbirds, swifts, and treeswifts), and shorter carpometacarpi (ostriches, tinamous, and rheas) relative to all other birds (Fig. 2A). Aequornithes (including loons, petrels, penguins, storks, herons, cormorants) occupied the majority of unique hind limb morphospace (0.012 log mm^3, 92.7% of unique volume) due to longer tibiotarsi in loons and the some procellariids such as the great shearwater (*Puffinus gravis*), as well as longer tibiotarsi and shorter tarsometarsi in penguins (Fig. 2B).

Locomotor disparity hypothesis. Overall limb disparity (Fig. 3A) and hind limb disparity increased significantly with locomotor disparity (Fig. S3D), whereas forelimb disparity was not significantly related to locomotor disparity (Fig. S3B) but instead increased significantly with species richness and clade age (Table 1). Evolutionary rates were not significantly related to locomotor disparity for either locomotor module (Figs. 3B, S3B,E; Table 1). Lineage densities were significantly lower in clades with higher locomotor disparity for both limbs analyzed together (Fig. 3C), but the relationship was not significant when limb modules were analyzed individually (Fig. S3, Table 1). Results were congruent using an alternative disparity metric that is less sensitive to sample size (mean pairwise distance, MPD), with clade age explaining a significant proportion of variation in forelimb disparity, and species richness and locomotor disparity each explaining a significant proportion of hind limb disparity (Table 1). The relationship between overall limb disparity and locomotor disparity was not significant using the MPD disparity metric, but instead clade age explained a significant proportion of variation in overall limb disparity (Table 1). These results were robust to an alternative phylogeny (Table S3) and removal of potential outlier taxa (Table S4). Eight clades had both larger hind limb morphospace and higher locomotor scores than the forelimb. Of these, Aequornithes ("waterbirds") had the highest locomotor score in both the forelimb (locomotor score = 6) and hind limb modules (locomotor score = 7; Fig. S3). Only one clade had both a higher forelimb locomotor score and forelimb morphospace size than in the hind limb: Strisores (forelimb locomotor score = 5, morphospace size = 0.014 log mm^3; Figs. S1-S3).

Tempo and mode of avian limb segment evolution. Temporal disparity analyses showed contrasting evolutionary patterns for the accumulation of disparity in the forelimb and hind limb (Fig. 4). Forelimb disparity was greater than expected under a Brownian motion model early on in avian evolution, whereas hind limb disparity did not differ from the pattern expected under Brownian motion (Fig. 4). Forelimb and hind limb disparity curves were similar after excluding paleognaths from the analysis (Fig. S7). Multivariate rate analyses showed significantly stronger covariation within the forelimb than within the hind limb or among limb modules (Fig. 5). Evolutionary rates were also higher in forelimb traits than in hind limb traits, with the exception of the tarsometatarsus, which was evolving faster than all other limb traits (Fig. 5). Rates of evolution were similar for stylopod elements (humerus and femur). Within limb evolutionary covariation was expectedly higher than between limb covariation, with one exception: tarsometatarsus-radius/ulna covariation was similar to femur-tibiotarsus covariation (Fig. 5). As expected, all limb measurements were significantly correlated with body mass (Fig. S5).

Discussion

Forelimb morphospace is objectively larger than hind limb morphospace

Different patterns of morphospace distribution were seen in the forelimb and hind limb of extant birds. Specifically, hind limb morphospace was roughly half the size of forelimb morphospace, but individual clades tended to occupy larger areas and overlapped more (Fig. 2; see Supplementary Methods). In contrast, forelimb morphospace showed a larger fraction occupied by single clades (Fig. 2). Unique regions of limb morphospace were dominated by taxa that display more specialized locomotor behaviors. For example, the largest areas of unique forelimb morphospace were occupied by taxa known for emphasizing aerial locomotion over terrestrial locomotion, including hummingbirds and swifts (Strisores clade), swallows (Passeriformes clade), and albatrosses (Aequornithes clade). These taxa clustered in three separate regions: these Strisores possess long carpometacarpi and short humeri, swallows show long radii/ulnae and short humeri, and albatross display short carpometacarpi and intermediate humeri and radii/ulnae (Fig. 2A). Each of these taxa exemplifies notably different flight styles, namely ‘hummingbird-style’ flight in hummingbirds, continuous flapping in swifts and swallows, and dynamic soaring in albatross (Bruderer et al. 2010; Del Hoyo et al. 1996-2016). Unique forelimb space in Strisores and swallows was linked to maneuverability (Middleton and Gatesy 2000). For Strisores, this may be due to their unique flight styles and increased area for attachment of the primary remiges (Saville 1950; Middleton and Gatesy 2000). In the case of albatrosses, elongate proportions in the humerus may be due to stability required in dynamic soaring (Middleton and Gatesy 2000).

Flightless paleognaths (ratites such as ostrich and emu) and flightless wing-propelled divers such as penguins and the great auk (*Pinguinus impennis*) also occupied unique forelimb morphospace, owing to proportionally longer humeri than other taxa. These two areas are not, however, overlapping: flightless paleognaths possess shorter distal wing segments, whereas wing-propelled divers have shorter radii/ulnae (Fig. 2A; also see Middleton and Gatesy 2000; Wang and Clarke 2014). Shortened distal wing elements are a common feature of flightless taxa, although paleognaths take this trend to an extreme (Livezey 1988; 1989; Middleton and Gatesy 2000; Livezey 2003). Proportionally shortened intermediate wing elements appear linked to efficient wing-propelled diving in volant and flightless divers (Livezey 1989; Middleton and Gatesy 2000). In penguins, the humerus makes up proportionally more of the forelimb than in volant wing-propelled divers, potentially due to flight loss. From a functional perspective, proportionally shorter distal wing elements aid in moving wings through water, a comparatively denser medium than air (Storer 1960). Waterbirds (clade Aequornithes) were previously shown to occupy a large area of unique forelimb morphospace (Wang and Clarke 2014). This unique morphospace is reduced in our sample to only include penguins and albatrosses, due to increased taxonomic sampling in Galloanserae (gamebirds), Mirandornithes (grebes and flamingos), Gruiformes (cranes, rails, and allies), Otidimorphae (bustards and cuckoos), and Charadriiformes (shorebirds such as gulls, auks, and coursers).

In the hind limb, Aequornithes dominated unique areas of morphospace due to long tibiotarsi (in loons) and long tibiotarsi paired with short tarsometatarsi (in penguins). Grebes, a convergent clade of specialist foot-propelled divers, approached loon morphospace but do not possess such elongate tibotarsi. Although not included in these analyses, extinct flightless foot-propelled diving hesperornithiform birds from the Cretaceous occupied unique hind limb morphospace typified by an extremely long tibiotarsus (Gatesy and Middleton 1997). More elongate tibiotarsi support extensive distal hind limb musculature in foot-propelled diving loons, grebes, and Hesperornithiformes, enabling efficient aquatic propulsion (Clifton et al. 2018). Penguins also occupy a unique region of hind limb morphospace characterized by an extremely abbreviated tarsometatarsus (Gatesy and Middleton 1997). Other taxa approach this space, consisting of taxa that primarily locomote using their wings, including fregatids, some hummingbirds and swifts, some parrots, and kingfishers (Gatesy and Middleton 1997). In contrast to penguins which are accomplished walkers, most of these taxa excluding some parrots rarely locomote on land, and some are nearly incapable of walking (Le Maho and Dewasmes 1984: Del Hoyo et al. 1996; 1997; 1999; 2001). As with foot-propelled divers and elongate tibiotarsi, the evolution of more abbreviate tarsometatarsi in wing-propelled divers and taxa that emphasize perching over walking indicates that certain locomotor behaviors might have consistent proportional signals, but may not be detected in analyses that examine all of avian diversity. Terrestrial and wading taxa show a similar effect: flamingos are waders that occupy unique morphospace due to elongate tarsometatarsi and short femora. Wading long-legged taxa such as stilts and avocets, the saddle-billed stork (*Ephippiorhynchus senegale*), as well as more terrestrial species including the secretary bird (*Sagittarius serpentarius*) and pratincoles, also approach this unique region of morphospace. Going forward, examining intra-clade evolution following the origin of novel locomotor behaviors might clarify this pattern, as well as other potential associations between limb disparity and behavior.

Limb morphospace and the acquisition of novel locomotor modes

We found that locomotor disparity and limb disparity are positively correlated across birds for both the combined (Fig. 3A) and hind limb data sets (Fig. S3D, Table 1). This pattern could result from either of two evolutionary processes: i) elevated rates of morphological evolution in clades with more locomotor modes or ii) acquisition of new locomotor modes enables exploration of novel regions of limb morphospace (e.g., Fig. 2B). Phylogenetic regression analyses suggest that rates of limb evolution are similar across clades with simple and complex locomotor behaviors (Figs. 3B, S3B,E, Table 1; all $P > 0.05$). However, we found that clades with more locomotor behaviors are less closely packed in morphospace, showing significantly lower lineage densities in overall limb morphospace ($P < 0.01$; Fig. 3C) and significantly higher mean pairwise distances in hind himb morphospace ($P = 0.03$; see Table 1). Together, these findings suggest that the evolution of new locomotor modes has allowed lineages to exploit new regions of hind limb morphospace (e.g., Baumgart et al. 2021; Falk et al. 2021; Orkney et al. 2021). By contrast, clade differences in forelimb disparity are primarily a function of clade age and species diversity (Table 1). These findings lend support to the hypothesis that locomotor modularity enabled more behavioral options for living birds. The recovered positive relationship between hind limb morphological disparity and locomotor disparity is intuitive, but interesting in light of the fact that consistent associations between limb morphometrics and ecology in birds have largely proved elusive (Shatkovska and Ghazali 2020), outside of some wading and diving taxa (Zeffer et al. 2003; Hinic-Frlog and Motani 2010; Stoessel et al. 2012; Wang and Clarke 2014). This may be due, in part, to differences in how morphological disparity relates to behavioral disparity between the forelimb and hind limb modules (Table 1), or that some previous studies assigned taxa to single ecological (e.g., arboreal, swimming, terrestrial) or flight style categories (e.g., continuous flapper, dynamic soarer; e.g., Baumgart et al. 2021). Instead, our approach codes taxa as a mosaic of multiple locomotor characteristics rather than singular classifications (see Taylor and Thomas 2014 for another example of this approach). Continued use of multivariate methods will continue to be a valuable strategy for testing hypothesized ecomorphological patterns (e.g., Mitchell and Makovicky 2014, Orkney et al. 2021) and studying morphological integration, a grand challenge in comparative vertebrate morphology (Danos et al. 2022). One caveat of our approach of counting clade-specific locomotor behaviors is that it treats all locomotor behaviors as functionally equivalent; for example, clades that used hovering and soaring would have the same locomotor score as a clade that uses wing propelled diving and sallying. Future analyses comparing limb morphological variation and locomotor behavior disparity could interrogate this relationship at the species-level in a phylogenetic context (e.g., using Markov models and multivariate discrete variables).

The multifunctional nature of avian limbs may be a major obstacle in recovering consistent associations between limb morphology and locomotor behavior (see Stoessel et al. 2012). Anatomical structure must address multiple functional demands, which can prevent the evolution of an optimum morphology for a particular function (Wainright et al. 2005). A role for multifunctionality and behavioral plasticity should also be considered when investigating linkages between avian morphology and locomotor behavior (see Diogo 2017 for a broader discussion). Phylogenetically-informed investigations of multiple character systems that unite locomotor behavior, ontogeny, and morphology, as well as studies of avian locomotor biomechanics and community-level ecomorphological trends in extant and extinct birds, are all needed to overcome these problems (e.g., Ricklefs and Travis 1980; Dial 2003a; Heers et al. 2011; Heers et al. 2014; Mitchell and Makovicky 2014; Smith and Clarke 2014; Xu et al. 2014; Ornkey et al. 2021; Danos et al. 2022).

Early gains of forelimb disparity in birds

Limb disparity accumulated differently between limbs during the evolution of birds (Fig. 4). Analyses of temporal trends in limb disparity supported a scenario wherein the majority of forelimb disparity evolved quickly among major avian clades early on in the avian radiation (Fig. 4). This pattern was driven primarily by divergence in forelimb morphology among flightless paleognath lineages (e.g., ostrich, emu, and rhea; see Figs. 2A, S7). These results are consistent with patterns recovered from smaller datasets and narrower taxonomic focus or sampling that suggest changes in the hand occur early in avian evolution (Nebreda et al. 2020) and that forelimb morphology does not often change significantly with the origin of novel flight styles (Fig. S3, Table 1; also see Wang and Clarke 2014; 2015). Yet, our results contrast with evolutionary patterns in limb morphospace previously reported from Mesozoic birds: stem avialans show similar forelimb disparity as non-volant theropods, whereas hind limb disparity diversified in stem pygostylians (Benson and Choiniere 2013). High phylogenetic signal across all traits (Table S1) is consistent with previous findings that phylogenetic effects must be taken into account when investigating the relationship between limb morphology and ecological or locomotor signals (e.g., Nudds et al. 2007; Stoessel et al. 2012; Wang and Clarke 2014; Baumgart et al. 2021; Orkney et al. 2021).

Support for independent avian limb modules

We find support for stronger limb integration within forelimb and hind limb modules than between them (Fig. 5). This result is consistent with a recent study that used a high-density morphometric dataset but in fewer taxa (Orkney et al. 2021). This partial limb module decoupling is consistent with the observed differences in disparity accumulation over time for the different limb modules (Fig. 4). Interestingly, stylopod-autopod (humerus-carpometacarpus, femur-tarsometatarsus) and zeugopod-autopod covariances (radius/ulna-carpometacarpus, tibiotarsus-tarsometatarsus) were similar for each limb (Fig. 5). This suggests potential developmental constraints, as limbs are homologous and involve similar genes/expression profiles (Young and Halgrimsson 2005). Stylopod-zeugopod covariances were significantly different across limbs (humerus-radius/ulna, femur-tibiotarsus), as relative segment lengths in the forelimb were more strongly integrated (Fig. 5). Forelimb traits showed significantly higher evolutionary rates than most hind limb traits, with the exception of the tarsometatarsus (Fig. 5). In the hind limb, the stylopod and zeugopod elements (femur and tibiotarsus) evolved at the same rate, whereas for the forelimb the zeugopod (radius/ulna) evolved significantly faster than the humerus (Fig. 5) and at the same rate as the autopods (carpometacarpals). The fastest evolving limb element was the hind limb autopod (tarsometatarsus). This different in rates between autopods might be reflective of stronger stabilizing selection on wing/forelimb morphology due to flight requirements, whereas the tarsometatarsus is able to evolve more freely possibly due to relaxed selection (Lahti et al. 2009). Taken together, our finding of weaker evolutionary covariation among forelimb and hind limb modules than within them is consistent with recent work using geometric morphometrics (Orkney et al. 2021) and work by previous authors who found no support for morphological integration between serial limb homologues (e.g., humerus and femur) in a sample of 174 specimens of 7 bird species (Bell et al. 2011). These results support the hypothesis put forward by Gatesy and Dial (1996) that limb modularity may have enabled differential evolutionary elaboration among modules, facilitating morphological and locomotor disparity in birds.

Conclusions

Extant avian clades that display higher disparity in locomotor behaviors show greater limb disparity. Furthermore, evolutionary trends in disparity differed between the forelimb and hind limb. Forelimb disparity evolved rapidly among clades early in avian evolution, whereas hind limb disparity diversified within rather than among clades. Phylogenetic effects are a significant factor in explaining the variation observed in avian limb morphology; comparative study of avian limbs should incorporate phylogeny. Avian forelimbs and hind limbs show different evolutionary rates, and evolutionary changes in one locomotor module do not closely correspond to predictable changes in the other module when considered across birds as a whole. These findings, in concert, support the hypothesis that the origin of separate locomotor modules aided the evolution of avian locomotor disparity, and that morphological disparity accumulated differently between modules. Although locomotor modularity may be linked to the evolution of diverse morphologies and, by extension, locomotor behaviors in birds, comparative examinations of morphological evolution following the origin of specific novel locomotor behaviors (e.g., wing-propelled diving, scansoriality) are necessary to understand how limb form and function evolved across different avian lineages.

References

Allen, V., Bates, K., Li, Z., and J. R. Hutchinson. 2013. Linking the evolution of body shape and locomotor biomechanics in bird-line archosaurs. Nature. 497(7447):104–107.

Alström, P., Barnes, K. N., Olsson, U. et al. 2013. Multilocus phylogeny of the avian family Alaudidae (larks) reveals complex morphological evolution, non-monophyletic genera and hidden species diversity. Mol. Phylogenet. Evol. 69(3):1043-1056.

Baumgart, S. L., P. C. Sereno, and M. W. Westneat. 2021. Wing Shape in Waterbirds: Morphometric Patterns Associated with Behavior, Habitat, Migration, and Phylogenetic Convergence. Integr Org Biol 3:obab011.

Bell, E., Andres, B. and A. Goswami. 2011. Integration and dissociation of limb elements in flying vertebrates: A comparison of pterosaurs, birds and bats. J. Evolution. Biol. 24(12):2586-2599.

Benson, R. B. J. and J. N. Choiniere. 2013. Rates of dinosaur limb evolution provide evidence for exceptional radiation in Mesozoic birds. P. R. Soc. B. 280:20131780.

Blomberg, S. P., Garland, T. and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution. 57(4):717–745.

Bruderer, B. Peter, D., Boldt A., and F. Liechti. 2010. Wing-beat characteristics of birds recorded with tracking radar and cine camera. Ibis. 152:272-291.

Burleigh, J. G., Kimball, R. T. and E.L. Braun. 2015. Building the avian tree of life using a large-scale, sparse supermatrix. Mol. Phylogenet. Evol. 84:53–63.

Carrano, M. T., and C. A.Sidor, C. A. 1999). Theropod hind limb disparity revisited: Comments on Gatesy and Middleton (1997). J Vertebr Paleontol. 19(3):602-605.

Clarke, J. A. and K. M. Middleton. 2008. Mosaicism, modules, and the evolution of birds: Results from a Bayesian approach to the study of morphological evolution using discrete character data. Syst. Biol. 57(2):185–201.

Clarke, J. A., Zhou, Z. and F. Zhang. 2006. Insight into the evolution of avian flight from a new clade of Early Cretaceous ornithurines from China and the morphology of *Yixianornis grabaui*. J. Anat. 208(3):287–308.

Clavel, J., G. Escarguel, and G. Merceron. 2015. mvMORPH: an rpackage for fitting multivariate evolutionary models to morphometric data. Methods Ecol. Evol. 6:1311–1319.

Clavel, J., Aristide, L., and H. Morlon. 2018. A penalized likelihood framework for high-dimensional phylogenetic comparative methods and an application to new-world monkeys. Syst. Biol. https://doi.org/10.1093/sysbio/syy045

Clifton, G. T., Carr, J. A., and A. A. Biewener. 2018. Comparative hind limb myology of foot-propelled swimming birds. J. Anat. 232(1):105-123.

Cooney, C. R., J. A. Bright, E. J. R. Capp, A. M. Chira, E. C. Hughes, C. J. A. Moody, L. O. Nouri, Z. K. Varley, and G. H. Thomas. 2017. Mega-evolutionary dynamics of the adaptive radiation of birds. Nature 542:344–347.

Danos, N., K. L. Staab, and L. B. Whitenack. 2022. The Core Concepts, Competencies, and Grand Challenges of Comparative Vertebrate Anatomy and Morphology. Integr Org Biol 4:obac019.

Dececchi, T. A. and H. C. E. Larsson. 2013. Body and limb size dissociation at the origin of birds: Uncoupling allometric constraints across a macroevolutionary transition. Evolution. 67(9):2741–2752.

del Hoyo, J., Elliot, A., J. Sargatal, et al. 1992-2016. Handbook of the Birds of the World. 1-16. Lynx Edicions. Barcelona, Spain.

Dial, K. P. 2003a. Evolution of avian locomotion: Correlates of flight style, locomotor modules, nesting biology, body size, development, and the origin of flapping flight. Auk. 120(4):941-952.

Dial, K. P. 2003b. Wing-assisted incline running and the evolution of flight. Science. 299(5605):402-404.

Dial, K. P., Heers, A. M., and T. R. Dial 2015. Ontogenetic and evolutionary transformations: Ecological significance of rudimentary structures. In Great Transformations in Vertebrate Evolution. 283-301. Dial, K. P., Shubin, N., and E. L. Brainerd (Eds.). The University of Chicago Press. Chicago, United States.

Diogo, R. 2017. Etho-eco-morphological mismatches, an overlooked phenomenon in ecology, evolution and evo-devo that supports ONCE (Organic Nonoptimal Constrained Evolution) and the key evolutionary role of organismal behavior. Front. Ecol. Evol. 5:3-20.

Doube, M., Yen, S. C. W., Kłosowski, M. M. et al. 2012. Whole-bone scaling of the avian pelvic limb. J. Anat. 221(1):21–29.

Dufort, M. J. 2016. An augmented supermatrix phylogeny of the avian family Picidae reveals uncertainty deep in the family tree. Mol. Phylogenet. Evol. 94:313-326.

Ericson, P. G., Anderson, C. L., Britton, T. et al. 2006. Diversification of Neoaves: Integration of molecular sequence data and fossils. Biol. Lett. 2(4):543-547.

Falk, A. R., J. C. Lamsdell, and E. Gong. 2021. Principal component analysis of avian hind limb and foot morphometrics and the relationship between ecology and phylogeny. Paleobiology 47:314–336. Cambridge University Press.

Foote, M. 1993. Contributions of Individual Taxa to Overall Morphological Disparity. Paleobiology 19:403–419.

Garcia-Ramirez, J. C., Gibb, G. C. and S. A. Trewick. (2014). Deep global evolutionary radiation in birds: Diversification and trait evolution in the cosmopolitan bird family Rallidae. Mol. Phylogenet. Evol. 81:96-108.

Gatesy, S. M. 1991. Hind limb scaling in birds and other theropods: Implications for terrestrial locomotion. J. Morphol. 209(1):83-96.

Gatesy, S. M. 1995. Functional evolution of the hind limb and tail from basal theropods to birds. In Functional Morphology in Vertebrate Paleontology. 219-234. Thomason, J.J. (Ed.). Cambridge University Press. Cambridge, England.

Gatesy, S. M. 2002. Locomotor evolution on the line to modern birds. In Mesozoic Birds: Above the Heads of Dinosaurs. 432-447. Chiappe, L. M., and L. M. Witmer (Eds.). University of California Press. Berkeley, Los Angeles, United States. London, England.

Gatesy, S. M. and K. P. Dial. 1996. Locomotor modules and the evolution of avian flight. Evolution. 50(1):331–340.

Gatesy, S. M. and K. M. Middleton. 1997. Bipedalism, flight, and the evolution of theropod locomotor diversity. J. Vertebr. Paleontol. 17(2):308–329.

Grafen, A. 1989. The phylogenetic regression. Philos. T. R. Soc. Lond. B. 326(1233):119-157.

Hadfield J. D. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. Journal of Statistical Software. 33(2):1-22. URL http://www.jstatsoft.org/v33/i02/.

Hamilton, N. 2016. ggtern: An extension to 'ggplot2', for the creation of ternary diagrams. R package version 2.2.0. URL: https://CRAN.R-project.org/package=ggtern.

Harmon, L. J., Schulte II, J. A., Larson, A. et al. 2003. Tempo and mode of evolutionary radiation in iguanian lizards. Science. 301:961-964

Harmon L. J., Weir, J. T. Brock, C. D. et al. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics. 24:129-131.

Harmon, L. J., Losos, J. B., Davies, T. J. et al. 2010. Early bursts of body size and shape evolution are rare in comparative data. Evolution. 64(8):2385-2396.

Heers, A. M., Tobalske, B. W. and K. P. Dial. 2011. Ontogeny of lift and drag production in ground birds. J. Exp. Biol. 214(5):717-725.

Heers, A.M., Dial, K. P. and B. W. Tobalske. 2014. From baby birds to feathered dinosaurs: Incipient wings and the evolution of flight. Paleobiology. 40(3):459-476.

Heers, A. M. and K. P. Dial. 2015. Wings versus legs in the avian bauplan: Development and evolution of alternative locomotor strategies. Evolution. 69(2):305–320.

Henningsson, P., Spedding, G. R. and A. Hedenström. 2008. Vortex wake and flight kinematics of a swift in cruising flight in a wind tunnel. J. Exp. Biol. 211(5): 717-730.

Hertel, F., Maldonado, J. E. and D. Sustaita. 2015. Wing and hind limb myology of vultures and raptors (Accipitriformes) in relation to locomotion and foraging. Acta Zool.-Stockholm. 96(3):283-295.

Hinić-Frlog, S. and R. Motani. 2010. Relationship between osteology and aquatic locomotion in birds: Determining modes of locomotion in extinct Ornithurae. J. Evolution. Biol. 23(2):372–385.

Ho, L. s. T. and C. Ané. 2014. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. Syst. Biol. 63(1):397-408.

Hutchinson, J. R. 2001a. The evolution of pelvic osteology and soft tissues on the line to extant birds (Neornithes). Zool. J. Linn. Soc.-Lond. 131(2):123–168.

Hutchinson, J. R. 2001b. The evolution of femoral osteology and soft tissues on the line to extant birds (Neornithes). Zool. J. Linn. Soc.-Lond. 131(2):169-197.

Hutchinson, J. R. 2002. The evolution of hind limb tendons and muscles on the line to crown-group birds. Comp. Biochem. Phys. A. 133(4):1051–1086.

Hutchinson, J. R. and V. Allen. 2009. The evolutionary continuum of limb function from early theropods to birds. Naturwissenschaften. 96(4): 423-448.

Ives, A. R. 2019. R^2s for Correlated Data: Phylogenetic Models, LMMs, and GLMMs. Syst. Biol. 68:234–251.

Jarvis, E. D., Mirarab, S., Aberer, A. J. et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science. 346(6215):1320–1331.

Karoullas, C., and R. L. Nudds. 2021. The link between avian brachial index, flight capability and the neornithine evolutionary radiation. J. Morphol. 282:1698–1707.

Kilbourne, B. M. 2013. On birds: Scale effects in the neognath hind limb and differences in the gross morphology of wings and hind limbs. Biol. J. Linn. Soc. 110(1):14–31.

Kimball, R. T., C. H. Oliveros, N. Wang, N. D. White, F. K. Barker, D. J. Field, D. T. Ksepka, R. T. Chesser, R. G. Moyle, M. J. Braun, R. T. Brumfield, B. C. Faircloth, B. T. Smith, and E. L. Braun. 2019. A Phylogenomic Supertree of Birds. Diversity 11:109. Multidisciplinary Digital Publishing Institute.

Lahti, D. C., Johnson, N. A., Ajie, B. C., Otto, S. P., Hendry, A. P., Blumstein, D. T., Coss, R. G., Donohue, K., and Foster, S. A. 2009. Relaxed selection in the wild. Trends In Ecology & Evolution. 24(9):487–496.

Le Maho, Y. and G. Dewasmes 1984. Energetics of Walking in Penguins. In Seabird Energetics. 235-244. Whittow, G. C. and I. H. Rahn (Eds.). Plenum Press. New York, United States.

Livezey, B. C. 1988. Morphometrics of flightlessness in the Alcidae. Auk. 105(4):681-698.

Livezey, B. C. 1989. Morphometric patterns in recent and fossil penguins (Aves, Sphenisciformes). J. Zool. 219(2):269-307.

Livezey, B. C. 2003. Evolution of flightlessness in rails (Gruiformes: Rallidae): Phylogenetic, ecomorphological, and ontogenetic perspectives. Ornithological Monographs. 53:1-654

Maia, R., Eliason, C. M., Bitton, P.-P., Doucet, S. M., & Shawkey, M. D. 2013. pavo: An R package for the analysis, visualization and organization of spectral data. Methods in Ecology and Evolution. 4: 906–913.

Middleton, K. M. and S. M. Gatesy. 2000. Theropod forelimb design and evolution. Zool. J. Linn. Soc.-Lond. 128:149–187.

Miles, D. B. and R. E. Ricklefs. 1984. The correlation between ecology and morphology in deciduous forest passerine birds. Ecology. 65(5):1629-1640.

Miles, D. B., R. E. Ricklefs, and J. Travis. 1987. Concordance of Ecomorphological Relationships in Three Assemblages of Passerine Birds. Am. Nat. 129:347–364. The University of Chicago Press.

Mitchell, J. S. and P. J. Makovicky. 2014. Low ecological disparity in Early Cretaceous birds. P. R. Soc. B. 281(1787):20140608.

Mitsuo, S. 1955. On the relation between the relative size of bones and the habits or locomotions in the bird (Passers); 1. Indices of bones the limbs and the habits or the locomotions. Bulletin of the Faculty of Education of Shinshiu University. 6:71-103.

Morlon, H., Lewitus, E., Codamine, F. L. et al. 2016. RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. Methods Ecol. Evol. 7:589-597.

Moyle, R. G. 2004. Phylogenetics of barbets (Aves: Piciformes) based on nuclear and mitochondrial DNA sequence data. Mol. Phylogenet. Evol. 30(1):187-200.

Nebreda, S. M., G. Navalón, I. Menéndez, T. Sigurdsen, L. M. Chiappe, and J. Marugán-Lobón. 2020. Disparity and macroevolutionary transformation of the maniraptoran Manus. Bull. Am. Mus. Nat. Hist. 440:183–203.

Norberg, U. M. 1979. Morphology of the wings, legs, and tail of three coniferous forest tits, the Goldcrest, and the Treecreeper in relation to locomotor pattern and feeding station selection. Philos. T. R. Soc. Lond. B. 287(1019):131-165.

Nudds, R. L., Dyke, G. J., and J. M. V. Rayner 2007. Avian brachial index and wing kinematics: Putting movement back into bones. J. Zool. 272(2):218-226.

Nudds, R. L., Atterholt, J., Wang, X. et al. (2013). Locomotory abilities and habitat of the Cretaceous bird *Gansus yumenensis* inferred from limb length proportions. J. Evolution. Biol. 26(1):150–154.

Orkney, A., A. Bjarnason, B. C. Tronrud, and R. B. J. Benson. 2021. Patterns of skeletal integration in birds reveal that adaptation of element shapes enables coordinated evolution between anatomical modules. Nat Ecol Evol 5:1250–1258.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature. 401(6756):877-884.

Paradis E., Claude, J. and K. Strimmer 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 20:289-290.

Paterno, G. B., C. Penone, and G. D. A. Werner. 2018. sensiPhy : An r -package for sensitivity analysis in phylogenetic comparative methods. Methods Ecol. Evol. 9:1461–1467.

Pinheiro J., Bates D., DebRoy S. et al. and R Core Team. 2015. nlme: Linear and Nonlinear

Mixed Effects Models. R package version 3.1-121. URL: http://CRAN.R-project.org/package=nlme.

Provini, P., Tobalske, B. W., Crandell, K. E. et al. 2012. Transition from leg to wing forces during take-off in birds. J. Exp. Biol. 215(23):4115-4124.

Provini, P., Tobalske, B.W., Crandell, K. E. et al. 2014. Transition from wing to leg forces during landing in birds. J. Exp. Biol. 217(15):2659-2666.

Prum, R. O., Berv, J., Dornburg, A. et al. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature. 526(7574):569–573.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Rabosky, D. L., Slater, G. J., and M. E. Alfaro. Clade age and species richness are decoupled across the eukaryotic tree of life. PLoS Biol. https://doi.org/10.1371/journal.pbio.1001381

Ramirez, J. L., Miyaki, C. Y. and S. N. Del Lama. 2013. Molecular phylogeny of Threskiornithidae (Aves: Pelecaniformes) based on nuclear and mitochondrial DNA. Genet. Mol. Res. 12(3):2740-2750.

Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217-223.

Ricklefs, R. E. and J. Travis 1980. A morphological approach to the study of avian community organization. Auk. 97(2):321-338.

Riede, T., Eliason, C., Miller, E. H. et al. 2016. Coos, booms, and hoots: The evolution of closed-mouth vocal behavior in birds. Evolution. 70(8):1734-1746.

Savile, D. B. O. 1950. The flight mechanism of swifts and hummingbirds. Auk. 67(4):499-504.

Shatkovska, O. V., and M. Ghazali. 2020. Integration of skeletal traits in some passerines: impact (or the lack thereof) of body mass, phylogeny, diet and habitat. J. Anat. 236:274–287.

Smith, N. A. and J. A. Clarke 2014. Osteological histology of the Pan-Alcidae (Aves, Charadriiformes): Correlates of wing-propelled diving and flightlessness. The Anat. Rec. 297(2):188-199.

Stoessel, A., Kilbourne, B. M. and M. S. Fischer. 2013. Morphological integration versus ecological plasticity in the avian pelvic limb skeleton. J Morphol. 274(5):483–495.

Storer, R. W. 1960. Evolution in the diving birds. In Proceedings of the XII International Ornithological Congress. 2:694-707. Bergmann, G., Donner, K.O., and L. von Haartman (Eds.). Tilgmannin Kirjapaino. Helsinki, Finland.

Taylor, G., and A. Thomas. 2014. Chapter 7: Adaptation in avian wing design. In Evolutionary Biomechanics: Selection, Phylogeny, and Constraint. 105-122. Oxford University Press. Oxford, England.

Tobalske, B. W., Warrick, D. R., Clark, C. J. et al. 2007. Three-dimensional kinematics of hummingbird flight. J. Exp. Biol. 210(13):2368-2382.

Wainwright, P. C., Alfaro, M. E., Bolnick, D. I. et al. 2005. Many-to-one mapping of form to function: a general principle in organismal design? Integr. Comp. Biol. 45(2):256-262.

Wang, X. and J. A. Clarke. 2014. Phylogeny and forelimb disparity in waterbirds. Evolution.

68(10):2847–2860.

Wang, X. and J. A. Clarke. 2015. The evolution of avian wing shape and previously unrecognized trends in covert feathering. P. R. Soc. B. 282(1816):20151935.

Wickham, H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag. New York, United States. URL: http://ggplot2.org.

Wright, N. A., Steadman, D. W. and C. C. Witt. 2016. Predictable evolution toward flightlessness in volant island birds. P. Natl. Acad. Sci. USA. 201522931.

Xia, C., Liang, W., Carey, G. J. et al. 2016. Song characteristics of Oriental Cuckoo *Cuculus optatus* and Himalayan Cuckoo *Cuculus saturatus* and implications for distribution and taxonomy. Zool. Stud. 55(38). DOI: 10.6620/ZS.2016.55-38

Xu, X., Zhou, Z., Dudley, R., et al. 2014. An integrative approach to understanding bird origins. Science. 346(6215):1253293.

Young N. M. and Hallgrímsson B. 2005. Serial homology and the evolution of mammalian limb covariation structure. Evolution. 59(12):2691-704.

Zeffer, A., Johansson, L. C. and Å. Marmebro. 2003. Functional correlation between habitat use and leg morphology in birds (Aves). Biol. J. Linn. Soc. 79(3):461–484

TABLES

Table 1. Results of phylogenetic generalized least squares (PGLS) analyses testing the relationship between locomotor disparity and clade-specific limb morphological disparity and rate metrics in a phylomorphospace context.

Module	Response	Predictors	Coef. ± S.E.	*t*	*P*	Partial R^2	AICc *w*
Forelimb	Sum morph. br. len.	**ln richness**	**1.10 ± 0.22**	**5.03**	**<0.01**	**0.62**	0.85
		ln clade age	**2.65 ± 0.91**	**2.92**	**0.01**	**0.31**	
	Mean pairwise dist.	**ln clade age**	**1.70 ± 0.74**	**2.32**	**0.04**	**0.16**	0.98
	Lineage density	ln richness	-1.43 ± 0.68	-2.12	0.05	0.24	0.99
	Evolutionary rate	ln clade age	1.43 ± 0.84	1.69	0.11	0.15	0.98
Hind limb	Sum morph. br. len.	**locom. score**	**0.27 ± 0.13**	**2.16**	**0.05**	**0.23**	0.85
		ln richness	**0.74 ± 0.28**	**2.64**	**0.02**	**0.33**	
	Mean pairwise dist.	**locom. score**	**0.15 ± 0.04**	**3.73**	**<0.01**	**0.55**	0.85
		ln richness	**-0.56 ± 0.12**	**-4.59**	**<0.01**	**0.60**	
	Lineage density	ln richness	-0.69 ± 0.50	-1.37	0.19	0.11	0.99
Both	Sum morph. br. len.	**locom. score**	**0.17 ± 0.08**	**2.16**	**0.05**	**0.25**	0.50
		ln richness	0.60 ± 0.29	2.06	0.06	0.23	
		ln clade age	1.32 ± 0.89	1.49	0.16	0.10	
	Mean pairwise dist.	ln richness	-0.23 ± 0.13	-1.76	0.10	0.17	0.87
		ln clade age	**1.55 ± 0.54**	**2.85**	**0.01**	**0.30**	
	Lineage density	**locom. score**	**-0.35 ± 0.14**	**-2.44**	**0.03**	**0.38**	0.94
		ln clade age	-4.99 ± 2.87	-1.74	0.11	0.11	

For each model, the most parsimonious set of predictors are presented (significant predictors in bold). Variables were removed in a bidirectional stepwise fashion based on AIC values with phylostep (Ho and Ané 2014). Partial R^2 values calculated by comparing model with and without that predictor using R2.lik (Ives 2019). AICc weights show reduced model support (see Table S2 for model selection). Results were robust to outliers (i.e., all results for locomotor score that were significant remained significant after running influ_phylm, and vice versa for non-significant results; see Table S4). Time-calibrated phylogeny of Kimball et al. (2019) was used in analyses.

Figure Legends

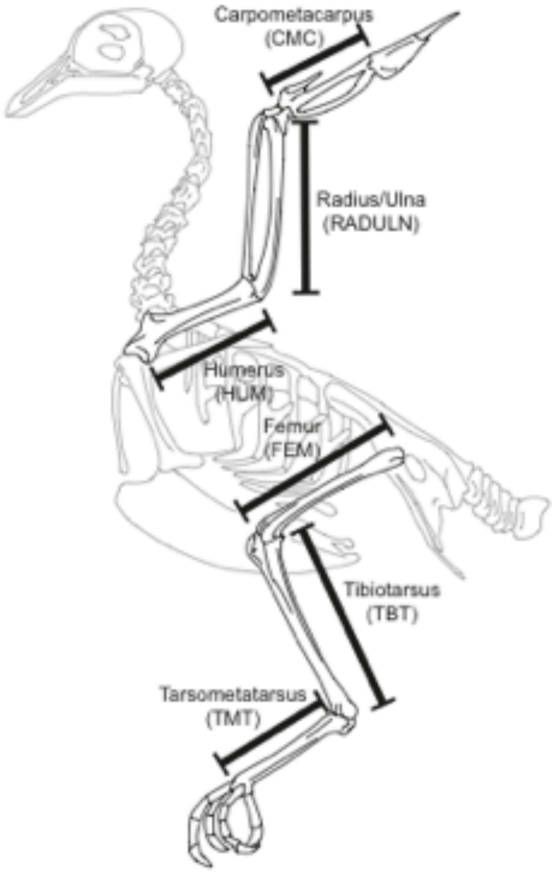

Figure 1. Diagrammatic representation of a bird skeleton indicating measurements for each limb. Total lengths were obtained for the humerus (HUM), radius/ulna (RADULN), and carpometacarpus (CMC) segments for the forelimb, and the femur (FEM), tibiotarsus (TBT), and tarsometatarsus (TMT) segments for the hind limb. Species visualized is the rock dove (*Columba livia*).

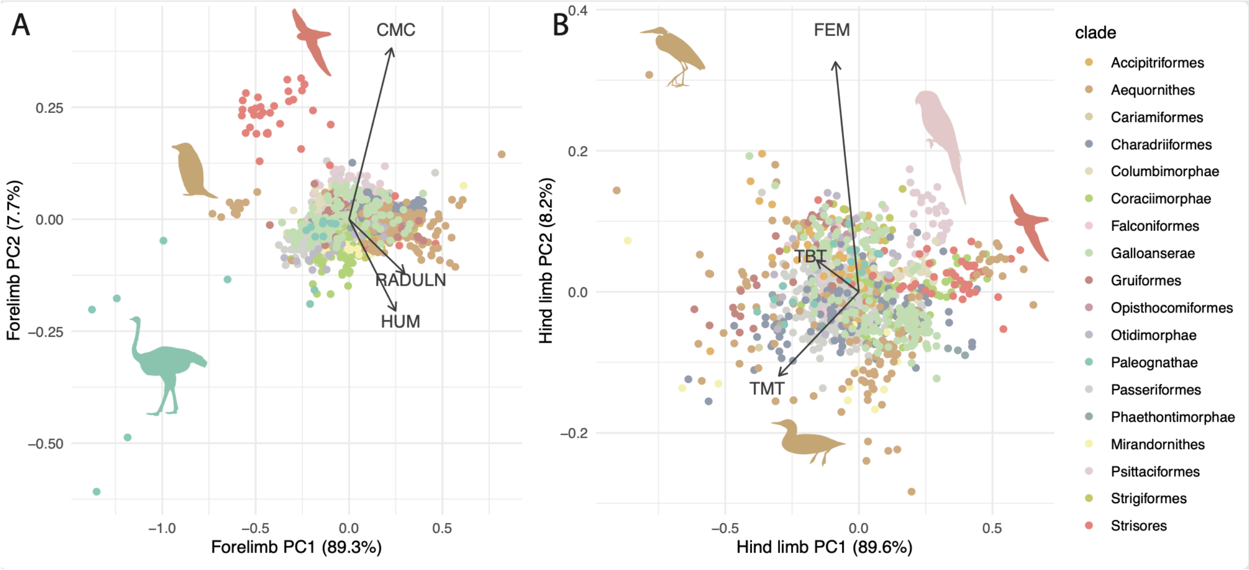

Figure 2. Avian limb morphospace. Principal component axes representing >95% variation in relative limb element lengths for the forelimb (A) and hind limb (B). Arrows represent PC loadings, with labels corresponding to limb segments (Humerus: HUM; Radius/Ulna: RADULN; Carpometacarpus: CMC; Femur: FEM; Tibiotarsus: TBT; Tarsometatarsus: TMT). Points represent mean values for species colored by clade. Silhouettes represent taxa that occupy unique positions in morphospace.

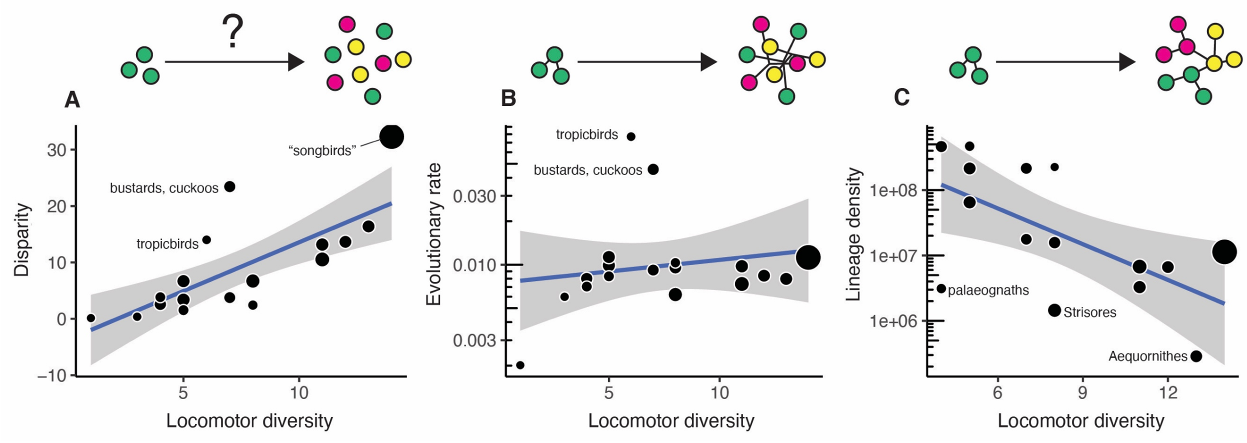

Figure 3. Correlated evolution of limb disparity and locomotor disparity in birds. (A) raw morphological disparity (sum of morphological branch lengths) increases with locomotor disparity (number of distinct locomotor modes in a clade; $P < 0.01$). (B) Evolutionary rates did not covary with locomotor disparity ($P = 89$). One example of a case where we would expect a positive relationship between these variables is if the evolution of novel locomotor behaviors leads to relaxed selection on limb morphology that translates into greater rates of limb evolution. (C) Lineage density was lower for clades with greater locomotor disparity ($P = 0.01$). This is expected if diverse clades partition morphospace through innovations associated with different locomotor (or foraging) modes. Size of points corresponds to log number of species in the clade. Outliers and focal clades labeled. Cartoons illustrate different scenarios, with color of circles representing novel locomotor modes and position of circles corresponding to a particular morphology. Shaded regions are linear model fits for visualization purposes only (see Table 1 for statistical output).

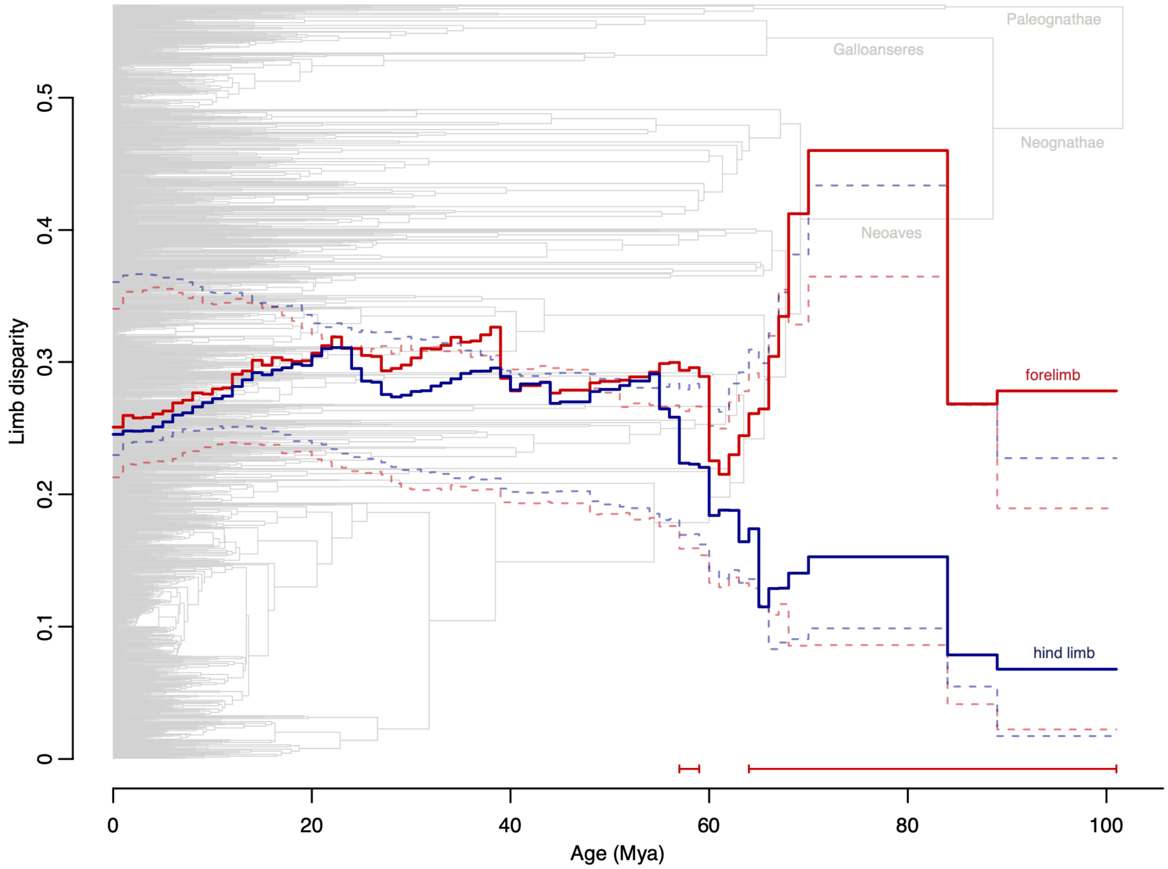

Figure 4. Accumulation of avian limb disparity over time. Solid lines show limb disparity calculated over different time bins for forelimb (red) and hind limb relative segment lengths (blue), based on the Burleigh et al. (2015) tree topology (N = 1140 species). The x-axis represents time, starting at the tips of the phylogeny and ending at the root. The y-axis represents mean disparity (pairwise Euclidean distances) calculated in 1-My time bins. Dashed lines represents 95% confidence intervals for null disparity curves estimated by 100 Brownian motion simulations. Segments at the bottom indicate regions of time over which forelimb disparity was significantly higher than hind limb disparity (see Methods for details).

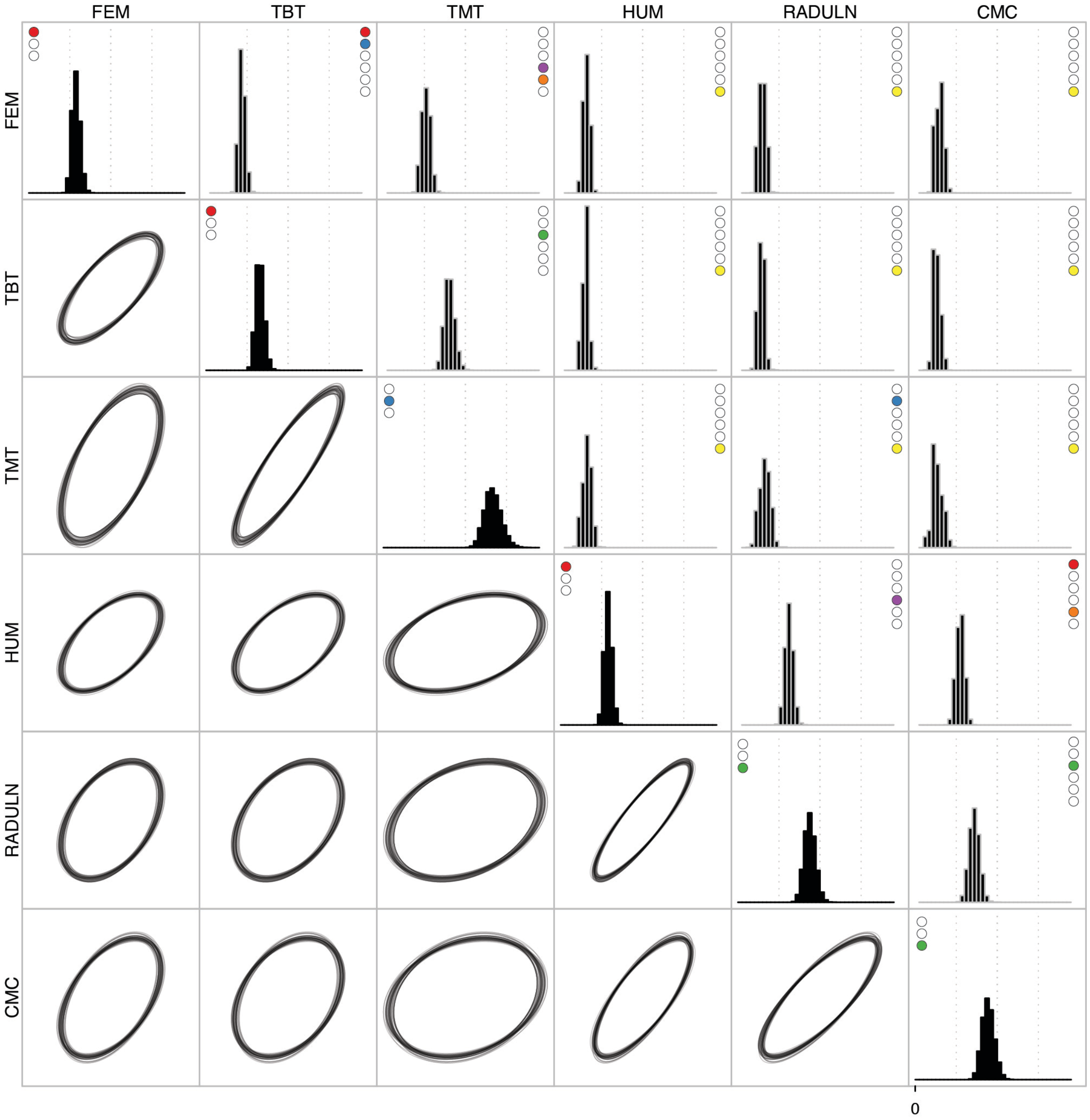

Figure 5. Evolutionary rates and covariation of relative limb segments in birds. Evolutionary rate matrix for three relative forelimb segment lengths: humerus (HUM), radius/ulna (RADULN), and carpometacarpus (CMC), and three hind limb segments: femur (FEM), tibiotarsus (TBT) and tarsometatarsus (TMT). Ellipses in lower off-diagonals display recovered

evolutionary relationships among traits using ratematrix. Histograms of Brownian evolutionary rates (σ^2) are shown on the diagonals (black histograms), and histograms of covariances are shown on the upper off-diagonals (gray histograms). Statistical comparisons were made between evolutionary rates and covariances by calculating Bayesian P values in MCMCglmm (Hadfield 2010), accounting for multiple tests with false discovery rate (FDR) P value adjustment. Small circles indicate significance of comparisons–any two rate (or covariance) histograms sharing similarly colored circles are not significantly different from each other (e.g., relative tarsometatarsus length is evolving significantly faster than femur length, and tibiotarsus-tarsometatarsus covariance is not significantly different from covariation between the radius/ulna and carpometacarpus). Note: only among-rate and among-covariance comparison can be made, as this is how we conducted significance testing.