doi:10.1111/evo.13893

Signal evolution and morphological complexity in hummingbirds (Aves: *Trochilidae*)

Chad M. Eliason,[1,2] Rafael Maia,[1] Juan L. Parra,[3] and Matthew D. Shawkey[4]

[1]*Grainger Bioinformatics Center, Field Museum of Natural History, Chicago*

[2]*E-mail: celiason@fieldmuseum.org*

[3]*Grupo de Ecología y Evolución de Vertebrados, Instituto de Biología, Universidad de Antioquia, Medellín, Colombia*

[4]*Evolution and Optics of Nanostructures Group, Department of Biology, University of Ghent 9000, Ghent, Belgium*

Received March 8, 2019

Accepted November 12, 2019

Understanding how animal signals are produced is critical for understanding their evolution because complexity and modularity in the underlying morphology can affect evolutionary patterns. Hummingbird feathers show some of the brightest and most iridescent colors in nature. These are produced by optically complex stacks of hollow, platelet-shaped organelles called melanosomes. Neither how these morphologies produce colors nor their evolution has been systematically studied. We first used nanoscale morphological measurements and optical modeling to identify the physical basis of color production in 34 hummingbird species. We found that, in general, the melanosome stacks function as multilayer reflectors, with platelet thickness and air space size explaining variation in hue (color) and saturation (color purity). Additionally, light rays reflected from the outer keratin surface interact with those reflected by small, superficial melanosomes to cause secondary reflectance peaks, primarily in short (blue) wavelengths. We then compared variation of both the morphological components and the colors they produce. The outer keratin cortex evolves independently and is more variable than other morphological traits, possibly due to functional constraints on melanosome packing. Intriguingly, shorter wavelength colors evolve faster than longer wavelength colors, perhaps due to developmental processes that enables greater lability of the shapes of small melanosomes. Together, these data indicate that increased structural complexity of feather tissues is associated with greater variation in morphology and iridescent coloration.

KEY WORDS: **Iridescence, macroevolution, melanosomes, ornaments.**

Sexual selection is thought to promote signal diversity (West-Eberhard 1983; Price 2008), but empirical support for this idea remains controversial (Parra 2010; Seddon et al. 2013; Huang and Rabosky 2014; Servedio and Burger 2014). This disagreement might be explained, in part, by a frequent lack of consideration of the mechanisms by which a signal is produced. Evolution of new signals involves a change in structure (whether it be chemical, morphological, or other) (Mayr 1960), thus, any study of signal diversity should consider the nature of the structure itself. Studies on evolutionary relationships between morphology and function are common in ecological traits (Alfaro et al. 2005; Stayton 2006; Wainwright 2007; Claverie and Patek 2013; Dumont et al. 2014), but less so for signal traits (Ord et al. 2013; Eliason et al. 2015). However, compared to classical examples of naturally selected traits, sexually selected signal traits are often under stronger directional selection (Hoekstra et al. 2001) and possess genetic correlations with mating preferences that could lead to runaway selection (West-Eberhard 1983; Price 2008; Prum 2010). This enhanced potential for variation among species makes signal traits an ideal system for studying whether and how the way signals are produced influences how they evolve.

Color is a major axis of phenotypic variation, particularly in birds, which produce diverse color signals through a combination of light absorption by pigments (pigment-based colors) and light scattering from nanostructured feather tissues composed of melanin, keratin, and air (structural colors) (Stoddard and

Prum 2011; Shawkey and D'Alba 2017). In general, birds with more complex nanostructures have more variable structural colors (Maia et al. 2013), and recent work in ducks (Aves: Anatidae) has shown how form-function innovations facilitate color evolution (Eliason et al. 2015). However, as birds produce structural colors in several ways (Durrer 1977), broader comparisons of structural color diversity are needed to shed light on the mechanisms and evolution of the shared system, and how sexual selection might shape morphological diversity.

Hummingbirds show explosive lineage diversification (McGuire et al. 2014) and display diverse locomotor (Clark et al. 2018), acoustic (Clark and Feo 2010), and visual signals (Parra 2010) in a variety of habitats and elevations. Bright iridescent plumage colors are present in nearly all hummingbird species (Greenewalt et al. 1960a,b; Parra 2010). A morphological innovation present in all hummingbird species examined to date is a flattened melanin granule (platelet) interspersed with air bubbles (Greenewalt et al. 1960a,b; Durrer 1977). Hollow melanosomes have evolved at least seven times across birds (Eliason et al. 2013), and flattening has evolved at least five times (Durrer 1977; Hu et al. 2018). However, copresence of these traits is less common, as these morphologies have, thus, far only been described in African starlings, trogons, and hummingbirds (Durrer 1977).

Stacks of hollow platelets in iridescent hummingbird feathers comprise some of the most complex nanostructures known in birds (Durrer 1977; Greenewalt et al. 1960a). Complex traits, whether morphological or behavioral (Leal and Losos 2015), are those traits requiring a greater number of subtraits (or parameters) to describe their variation. Vermeij (1973) laid out a framework for testing how increased complexity might influence the variability of complex morphological traits. In particular, three distinct aspects can be important for evolutionary variability: (1) the number of morphological subtraits, (2) the evolutionary independence among subtraits, and (3) the range (or evolutionary rates) of individual subtraits. In general, higher values in any of these aspects would lead to greater evolutionary variability. In the context of functional traits like the iridescent nanostructures of hummingbirds, it is necessary to understand not only the evolution of morphological subtraits but also the colorful signals they produce. Greenewalt et al. (1960a) was the first to study color mechanisms in hummingbirds using a combination of empirical work and theoretical (optical) calculations. Durrer (1977) recognized that hummingbird melanosomes have additional interfaces for light reflection (e.g., between melanin and air) that would make optical calculations more challenging than with simpler nanostructures. More recently, Giraldo et al. (2018) studied the color-producing mechanism of Anna's hummingbird (*Calypte anna*) using advanced optical simulations. Both Greenewalt et al. (1960) and Giraldo et al. (2018) described secondary reflectance peaks and implicated a superficial layer of keratin (the cortex) in the production of these peaks. Giraldo et al. (2018) further noted miniature platelets lying directly beneath the cortex. However, previous work on hummingbird iridescence has been primarily descriptive (Durrer 1977), based on small sample sizes (Giraldo et al. 2018; Greenewalt et al. 1960a), or relied on methods that infer morphology from spectral curves rather than by direct microscopic examination (Greenewalt et al. 1960a). These aspects make it difficult to answer questions about how diverse morphology-color relationships may have evolved. For example, what subtraits account for the variation in coloration among hummingbird species?

Here, we take a bottom-up approach by using morphological parameters measured from transmission electron microscope (TEM) images of 44 feathers from 34 hummingbird species to simulate theoretical reflectance spectra that we compare with empirical reflectance spectra. We then address the evolution of both morphology and iridescent coloration using multivariate comparative methods. Specifically, we ask (1) which morphological subtraits have the largest influence on the resulting phenotype (color), (2) to what extent subtraits are correlated with one another, and (3) to what extent do these relationships differ between two taxonomic groups differing in morphological complexity (i.e., hummingbirds and ducks). Our results have implications for understanding the origins and current state of functional diversity in iridescent coloration.

Materials and Methods

SAMPLING IRIDESCENT FEATHERS

We sampled 44 feathers from various body regions (31 gorget/throat feathers, 12 crown feathers, and 1 back/mantle feathers) in 34 hummingbird species (see Dataset S1 for sources). To assess whether our species sampling was uniform relative to lineages in a recent phylogeny (McGuire et al. 2014), we calculated two metrics of phylogenetic clustering: mean pairwise distance (MPD) and mean nearest taxon distance (MNTD), using the R package picante (Kembel et al. 2010). We then randomly sampled subsets of 34 species from the 293 recognized hummingbird species (McGuire et al. 2014) and calculated MPD and MNTD for 500 random subsets to create null distributions. Since either clustering or overdispersion of sampled species might be problematic when comparing evolutionary rates among clades, we calculated two-tailed P values as the proportion of null values less than (indicating clustering) or greater than (indicating overdispersion) the observed values. In both cases, we obtained P values > 0.05, suggesting our sampling was roughly uniform with respect to phylogeny (Fig. S1). To test the potential effects of sparse sampling on our comparative analyses, we ran 100 replicates of a jackknife procedure in which we randomly removed 33% of species at each iteration before refitting evolutionary models for nanostructural

and spectral traits, following the approach of Denton and Adams (2015).

MEASURING MORPHOLOGICAL TRAITS

We embedded 44 feather barbs in resin, cut cross-sections with a Leica UC-6 ultramicrotome (Leica Microsystems GmbH, Wetzlar, Germany), and imaged them with a TEM(JEOL JEM-1230) following Shawkey et al. (2003). We measured the following optically relevant traits for several images per specimen: platelet thickness (pt), platelet spacing (ps), air space diameter (air), keratin cortex thickness (cor), number of melanin platelets (layers), and top platelet thickness (pt_{top}; see Fig. 1B for schematic, see Dataset S2 for species means). In total, we took 4417 measurements from 135 images of individual feathers plucked from the crown ($N = 12$), gorget ($N = 31$), and back ($N = 1$) of 34 hummingbird species. For all morphological traits but layer number, we took 10 measurements haphazardly selected along the surface of a barbule. The number of platelet layers was observed to be ~uniform across the region imaged with the TEM. Therefore, a single value quantifies the overall nanostructure. Several (14/34, 40%) species have evolved small platelets at the tops of barbules (Figs. 1B, S2), previously only described in a single species (Giraldo et al. 2018), and we noted their presence or absence (Dataset S1). We examined repeatability between observers by plotting per-image measurement values for two different observers (M.D.S. and R.M.). There was a strong correlation between observers (intraclass correlation values ranged from 0.84-0.92), so we calculated averages from pooled measurements for each trait. To compare variation in morphological traits across different levels of organization (e.g., species, feathers), we fit linear mixed models using lmer in the lme4 package (Bates et al. 2015). Each morphological trait was used as the response, with TEM image, feather patch, and species included as random effects. This approach is similar to a nested taxonomic ANOVA (Starck and Ricklefs 1998) and allowed us to assess our sampling design (see Fig. S3). All morphological traits were natural log-transformed prior to analyses to achieve normality and to allow us to compare evolutionary rates among traits with different characteristic variabilities (discussed in Adams 2013).

QUANTIFYING COLOR

To understand whether and how color changes with angle, a hallmark of iridescent traits, we varied the light and viewer angles in tandem (i.e., in specular configuration) from 10° to 45° in 5° increments using a laboratory-made goniometer (Meadows et al. 2011) with an attached spectrophotometer and xenon light source (Avantes Inc., Broomfield, CO, USA). For some species, color was either black or too drab, and we could therefore not measure signal directionality. In total, we measured angle-resolved reflectance for 31 out of 44 feather specimens. For all spectra,

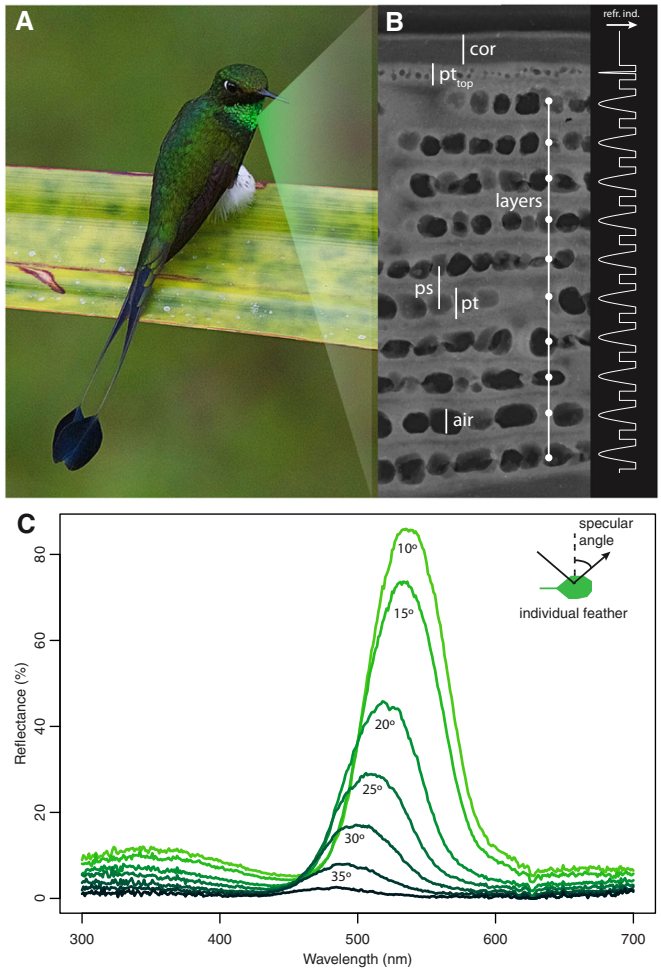

Figure 1. Measuring iridescent color and morphology in hummingbirds. (A) White-booted racket-tail hummingbird (*Ocreatus underwoodii*) with iridescent gorget feathers. (B) Transmission electron microscope (TEM) image of a cross-section of an iridescent barbule showing the six traits we measured: cortex thickness (*cor*), top platelet thickness (pt_{top}), number of platelets (*layers*), platelet spacing (*ps*), platelet thickness (*pt*), and air space diameter (*air*). (C) Measurement geometry (inset) and reflectance spectra at various specular angles used to confirm iridescence (i.e., color change with angle) across hummingbirds.

we calculated three peak shape variables for each spectrum: *hue* as the wavelength at which the reflectance reaches a maximum, *brightness* as height of the main peak, and *saturation* as the width of the peak at the midpoint between the maximum and minimum reflectance value. In addition to the primary peak (from which we calculated hue values), several species have secondary reflectance peaks occurring at shorter wavelengths. Importantly, these secondary peaks are not simply harmonics of the main peak (e.g., occurring at 1/2 the main peak's hue) and can therefore not be explained by standard interference models. To understand morphological predictors of secondary peaks, we noted their presence

or absence (e.g., see arrows in Fig. S6) for statistical analysis (see below).

PREDICTING SPECTRAL SHAPE FROM MORPHOLOGY

Iridescent nanostructures in birds are complex and involve the alignment of melanosomes in a single direction (often parallel to the axis of barbules). Scanning electron images of longitudinal sections in several hummingbird species (Hu et al. 2018) revealed that air spaces were roughly spherical and uniformly distributed throughout melanosomes (i.e., did not show any long-range order that might cause diffraction). Furthermore, previous results on polarization in the ruby-throated hummingbird showed no difference in reflectance at different polarizations (Eliason and Shawkey 2012). These observations suggest little ordering in two dimensions, so we used a one-dimensional optical model to simulate reflectance. This model essentially slices the nanostructure up into pieces with uniform refractive index and then uses a transfer matrix approach (Jellison 1993) to simulate light reflection at the interface of each layer. To define slices, we calculated refractive index as a function of position through an air sphere within a platelet as

$$n_{\text{avg}} = \Phi\,(n_{\text{mel}} - n_{\text{air}})\left[1 - (z/d_{\text{air}} + 1)^2\right]$$

where Φ is the density of air spheres, z is the vertical position through a sphere, $n_{\text{air}} = 1$, and d_{air} is the diameter of the air space (see Fig. 1B, outset) (Diamant et al. 2012). We generated wavelength-dependent complex refractive indices as $\tilde{n} = A + B/\lambda^2 - iC\exp(-\lambda/\lambda_i)$, where λ is the wavelength. We used published parameter values for melanin ($A = 1.648$, $B = 23{,}700$, $C = 0.56$, and $\lambda_i = 270$ nm) (Stavenga et al. 2015) and keratin ($A = 1.532$, $B = 5890$) (Leertouwer et al. 2011). For each layer, we computed volume-average refractive indices following Garahan et al. (2007). We tried various resolutions for how the structure was "sliced" into uniform layers and found that 100 layers gave similar results as 500 (Fig. S5), thus, we settled on a resolution of 100 for all simulations. We removed 6 of 44 samples prior to form-function analyses because microscopy revealed that feather barbules were either from a white or black (i.e., noniridescent) region of the feather or had disordered melanosome arrays not expected to produce iridescence. We simulated reflectance for the remaining 38 iridescent crown and gorget feathers under four distinct models: (1) a full model of the barbule nanostructure with all six morphological traits measured (see above), (2) a reduced model without including keratin between platelets, (3) a reduced model without including small platelets at the barbule surface, and (4) a reduced model without a keratin cortex.

To compare the relationship between measured and predicted spectra for the full (six-parameter) optical model, we used two-block partial least squares (PLS) implemented in the geomorph R package (Adams and Otárola-Castillo 2013). This method has been used widely in geometric morphometrics to study correlation in multivariate datasets (e.g., see Rohlf and Corti 2000). In our case, multivariate X and Y were the simulated and empirical reflectance values, respectively, both normalized to have maximum values of 1. We assessed significance of the test statistic (r_{PLS}) by randomly shuffling species in X relative to those in Y 999 times. To further compare different optical models, we calculated mean residual sum of squares (RSS) between normalized measured and predicted spectra for all four models, with lower values indicating a better fit.

LINKING MORPHOLOGICAL FORM TO OPTICAL FUNCTION

Statistical modeling

To test for statistical associations between color and morphological traits, we fit separate models for each peak shape variable (hue, saturation, and brightness) as the response and our six morphological traits as predictors. We fit full additive models with ordinary least squares (OLS) and removed nonsignificant terms ($P > 0.05$) in a backward step-wise fashion to determine the most parsimonious model for each color variable. We further assessed relative variable importance (i.e., the number of reduced models in which it is included) using the dredge and importance functions in the R package MuMIn v. 1.42.1. To account for phylogenetic signal in the data, we refit all reduced models in a phylogenetic mixed model framework implemented in MCMCglmm (Hadfield and Nakagawa 2010). To our knowledge, this is the only approach that accounts for phylogeny while also allowing for intraspecific variation (i.e., among patches).

To test for an association between the presence of secondary peaks and small, superficial platelets, we fit a phylogenetic generalized linear model with the phylolm R package (Ho and Ané 2014). To further verify the origin of secondary peaks, we determined the percentage of full optical simulations that correctly predicted secondary peaks to the accuracy of a reduced model not incorporating the upper cortex or small superficial platelets (see Dataset S3). We then used binomial tests to determine whether each model significantly predicted secondary peaks more often than expected at random.

Optical sensitivity analysis

To further assess the relative sensitivity of different morphological parameters on the shape of predicted reflectance spectra, we chose a species with a distinct secondary reflectance peak, *Ocreatus underwoodii* (see Fig. 1A), and simulated spectra while varying each parameter in turn and holding all other parameters constant at their mean values for this species. We varied: (1) platelet thickness ± 10 nm from the mean value (123-143 nm), (2) platelet spacing from the mean platelet thickness (i.e., with

platelets touching) to the observed spacing value (133-175 nm), (3) air space diameter from zero to the observed value (0-76 nm), (4) top platelet thickness from the observed value to the mean thickness of "regular" platelets (64-133 nm), (5) cortex thickness from zero to the observed value (0-154 nm), and (6) layer number ± 5 layers around the mean value (8-17 layers). All simulations were done in 10 equidistant steps.

TESTING THE BIOLOGICAL VERSATILITY HYPOTHESIS

Hummingbirds have one of the most complex nanostructures known in birds (Durrer 1977), varying in at least six dimensions (Figs. 1B, 3). To test Vermeij's (1973) "biological versatility" hypothesis that increases in complexity produce a wider range of potential forms (i.e., greater morphological disparity), we determined evolutionary correlations and variability among morphological subtraits and then compared these values between hummingbirds and dabbling ducks, a monophyletic clade for which careful studies have identified four morphological subtraits responsible for their variation in color (Eliason et al. 2015). We used a published dataset (Eliason et al. 2015) for 38 duck species, along with additional measurements of cortex thickness (Dataset S4). Since gorget feathers are most commonly used in mating displays and agonistic interactions in hummingbirds (Bleiweiss 1992), and because sample size for crown feathers was relatively low ($N = 11$), we used iridescent gorget feathers for our comparative analyses.

Morphological disparity measures the range of potential phenotypes a clade can occupy (Hughes et al. 2013) and is affected both by rates of and constraints on trait evolution (Felice et al. 2018). Under a Brownian motion (BM) model of trait evolution, species trait values "spread out" as a function of time, resulting in a monotonic increase in the amount of variation over time. By contrast, under an Ornstein-Uhlenbeck (OU) model, disparity at first builds up but then plateaus over time as variance-generating mechanisms (σ^2, represented as stochastic changes per unit time) are balanced by variance-reducing mechanisms (α, often described as a "rubber band" pulling phenotypes to some optimal, or constrained, value). The ratio between these two parameters ($\sigma^2/2\alpha$) is the stationary variance of the OU process (Bartoszek et al. 2012) and can be thought of as the steady-state amount of morphological disparity within a clade.

To estimate evolutionary covariance matrices, we used recent time-calibrated phylogenies (McGuire et al. 2014; Eliason et al. 2015) to fit three distinct models of multivariate trait evolution: (1) a Brownian motion model (BM), (2) a one-regime Ornstein-Uhlenbeck (OU) model in which trait evolution is constrained by a restraining parameter α, and (3) an early burst (EB) model in which trait evolution slows down toward the tips of the tree. We fit models using fit_t_pl (Clavel et al. 2018) in the RPANDA R package and selected among models using the generalized information criterion (GIC) following Clavel et al. (2018). Because evolutionary rates scale with α in OU models (Bartoszek et al. 2012), we estimated the stationary covariance matrix (see above) from the resulting rate matrices using custom R code (Eliason et al. 2015) and then calculated standard errors of parameters using published R code (Clavel et al. 2018). Evolutionary variances, taken as the diagonals of this transformed evolutionary covariance matrix, correspond to morphological disparity and are equivalent to coefficients of variation when using natural log-transformed data (see Gingerich 2009). Importantly, calculation of stationary covariance matrices removes the effect of time similarly for all traits (since we only calculate a single value of alpha), therefore, the relative widths of standard errors among traits remain the same under this transformation.

To compare evolutionary variation and flexibility in spectral shape between these clades, we estimated evolutionary covariance matrices using natural log-transformed raw reflectance values in 8-nm bins (Dataset S5). We then compared spectral disparity as above, and spectral flexibility (i.e., evolutionary independence among distinct wavelength bands) as Van Valen's mean coefficient of determination (i.e., the average of squared pairwise evolutionary correlations) (Van Valen 1974) derived from the estimated evolutionary covariance matrix. This is an overall measure of integration within a set of traits, with lower values indicating greater flexibility in spectral shape.

Results

GENERAL OBSERVATIONS

All barbules contained stacks of melanin platelets (varying in number from 2-16 layers) with roughly spherical air holes (Fig. 1B). Surrounding these stacks was a thin (10-207 nm) keratin cortex. Small platelets were observed just below the cortex of barbules in 16 species (see Figs. 1B, S2). Of these species, 11 had small air-filled platelets and five species had small solid (e.g., Fig. S2C) platelets (Dataset S1). In all cases, reflectance spectra changed considerably with angle (Fig. S4), confirming the structural origin of bright colors across hummingbirds.

DOES MORPHOLOGY PREDICT SPECTRAL SHAPE?

Overall, optical simulations incorporating all six morphological parameters (cortex thickness, platelet thickness, air space diameter, top platelet thickness, platelet spacing, and number of platelets) fit the observed spectra fairly well ($r_{PLS} = 0.559$, $P = 0.013$; Fig. S6). A model not including space between platelets slightly outperformed the full, six-parameter model (Figs. S7, S8B). Simulations excluding the outer keratin cortex and small superficial platelets showed a poorer fit (higher RSS values) relative to the full and "no spacing" simulations (Fig. S8B). Both the

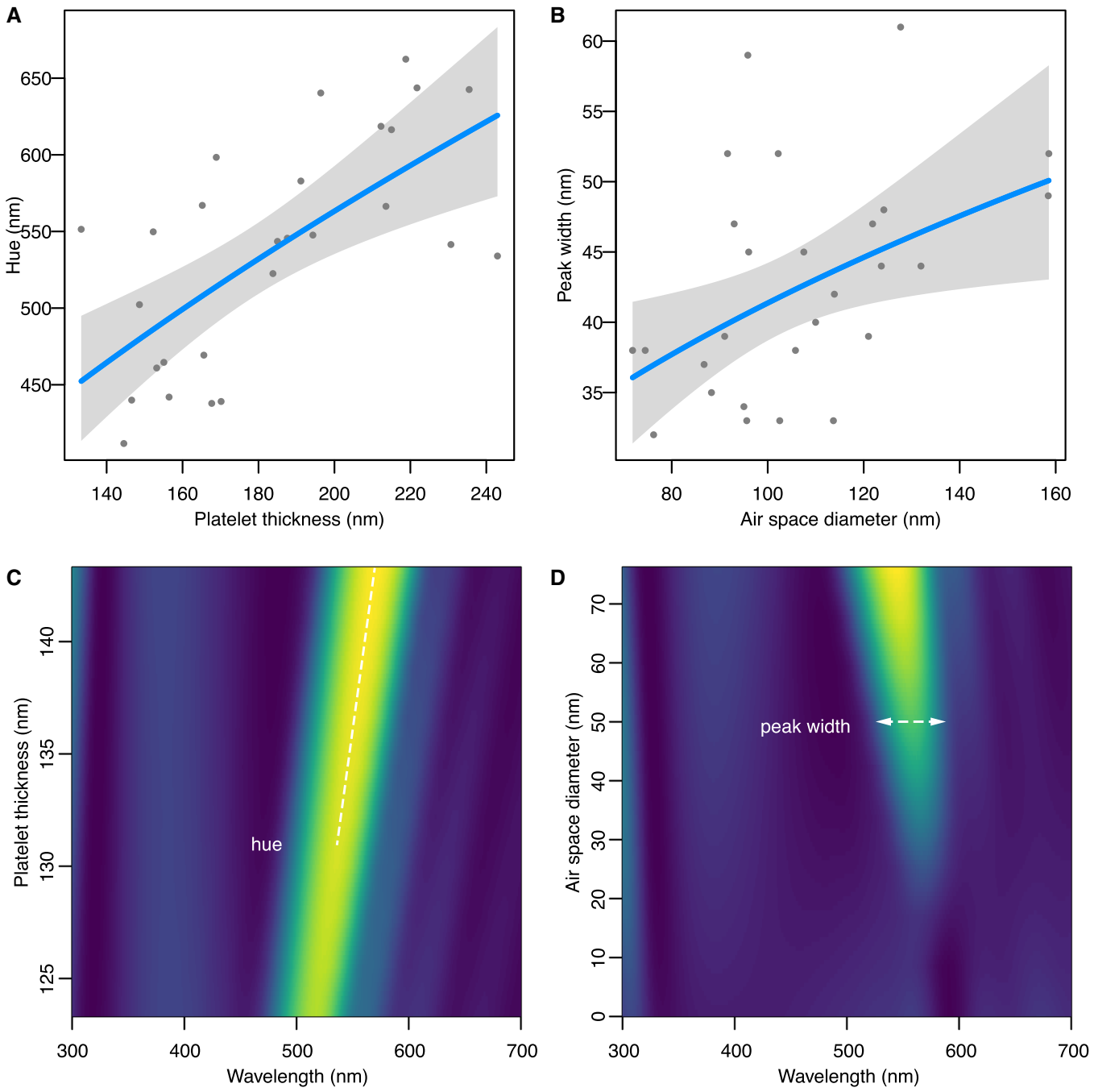

Figure 2. **Morphology-color relationships in hummingbirds. (A,B) Regression results showing the relationship between hue and platelet thickness (A) and between peak width and air space diameter within melanosomes (B). Points are partial residuals derived from the best-fitting models (see Table 1). (C,D) Optical simulations based on the morphology of *Ocreatus underwoodii*. Colors correspond to spectral reflectance, ranging from low (blue) to high (yellow). Simulated hues increase with platelet thickness (C) and reflectance peak widths broaden with the amount of air in melanosomes (D).**

full and no-spacing models captured variation in the width of the primary peak, as well as the location (hue, in nm) of the secondary peak (Figs. S6, S7). For 10 of 38 feathers there was a considerable mismatch between simulated and measured reflectance spectra (e.g., gorget feathers of *Coeligena bonapartei*; Fig. S6). Variation in platelet thickness primarily affects hue (Fig. 2C). Variation in cortex thickness affects peak width and, in some cases, can cause destructive interference (Yoshioka et al. 2012) giving "bimodal" reflectance peaks (e.g., for thicknesses ~40 nm; Fig. S9C). Both the thickness of the top platelet and spacing between adjacent

Table 1. Multiple linear regression model results for color variables.

Response	Predictor	Full OLS	Reduced OLS	Reduced PGLS	Importance
Hue	Spacing	−0.15 ± 0.21	...	...	0.25
	Air	−0.13 ± 0.14	...	...	0.24
	Layers	0.12 ± 0.10	...	...	0.42
	Platelet	**0.76 ± 0.22**	**0.54 ± 0.13**	**0.56 [0.24, 0.90]**	**0.99**
	Cortex	−0.04 ± 0.03	−0.05 ± 0.03	−0.05 [−0.12, 0.02]	0.49
	$F =$	5.09	11.27	...	...
	$df =$	5, 21	2, 24	...	...
Saturation	Spacing	0.05 ± 0.34	...	...	0.23
	Air	**0.33 ± 0.23**	**0.42 ± 0.16**	**0.37 [0.00, 0.71]**	**0.73**
	Layers	0.02 ± 0.16	...	...	0.21
	Platelet	0.15 ± 0.36	...	...	0.34
	Cortex	0.00 ± 0.05	...	...	0.19
	$F =$	1.29	6.77	...	...
	$df =$	5, 21	1, 25	...	...
Brightness	Spacing	−0.85 ± 1.45	...	...	0.22
	Air	0.33 ± 0.99	...	...	0.29
	Layers	0.12 ± 0.70	...	...	0.23
	Platelet	1.46 ± 1.56	...	...	0.35
	Cortex	−0.14 ± 0.23	...	...	0.24
	$F =$	0.52	...	...	...
	$df =$	5, 21	...	...	...

For each color variable, the full additive model is shown, with significant ($P < 0.05$) predictors for the best-fitting (reduced) model highlighted in bold. Models were fit with both ordinary least squares (OLS) and phylogenetic least squares (PGLS). Values are β ± SE (for OLS models) or Bayesian credible intervals in square brackets (for PGLS models). Relative variable importance calculated as the sum of Akaike weights for all models in which a predictor is included (see Methods).

platelets affected the brightness of the main peak (Fig. S9A, B), while spacing had a much stronger effect on hue (Fig. S9A). Peak width narrows with increasing number of layers (Fig. S9D).

WHAT MORPHOLOGICAL TRAITS BEST PREDICT COLOR VARIABILITY?

The best model for hue included platelet thickness and cortex thickness as predictors (Table 1). Hue significantly increased with platelet thickness (Fig. 2A, Table 1). For peak width, the best model included only air space diameter as a predictor (Table 1). Peak width increased significantly with the size of air spaces within platelets (Fig. 2B, Table 1). There were no significant predictors for brightness. These results were qualitatively similar when accounting for phylogeny using PGLS (see Table 1).

Small platelets were significantly associated with secondary reflectance peaks (phylogenetic GLM, $P = 0.041$). Optical simulations showed that secondary reflectance peaks can be predicted from optical theory (Fig. 4B). Specifically, secondary peaks depend on the interaction between the thicknesses of the keratin cortex and the top layer of small platelets. An optical model accounting for small platelets and cortex thickness correctly predicted 81.0% (binomial test, $P = 0.0036$) of secondary peaks in the empirical data, while a model not accounting for these traits predicted secondary peaks only 14.3% of the time ($P = 1$).

DOES COMPLEXITY INCREASE MORPHOLOGICAL DISPARITY?

In all cases, the OU model performed best, suggesting low phylogenetic signal for the traits we measured (Table S1). The cortex was more variable than, and was decoupled from, other morphological traits within hummingbird gorget feathers (Fig. 3B). By contrast, cortex thickness evolved in tandem with size and spacing of melanin granules in ducks (Fig. 3A). The thickness of the top layer of platelets was significantly more variable than the size of nonmodified platelets (Figs. 3B, S10). The number of platelet layers decreased significantly with spacing (Fig. 3B). For spectral data, evolutionary disparities in hummingbirds were elevated primarily in blue (400-450 nm) and yellow-red wavelengths (600-700 nm; see Fig. 4C). Hummingbirds showed greater evolutionary flexibility in spectral shape compared to ducks (Van Valen's mean $R^2 = 0.35$ and 0.45, respectively).

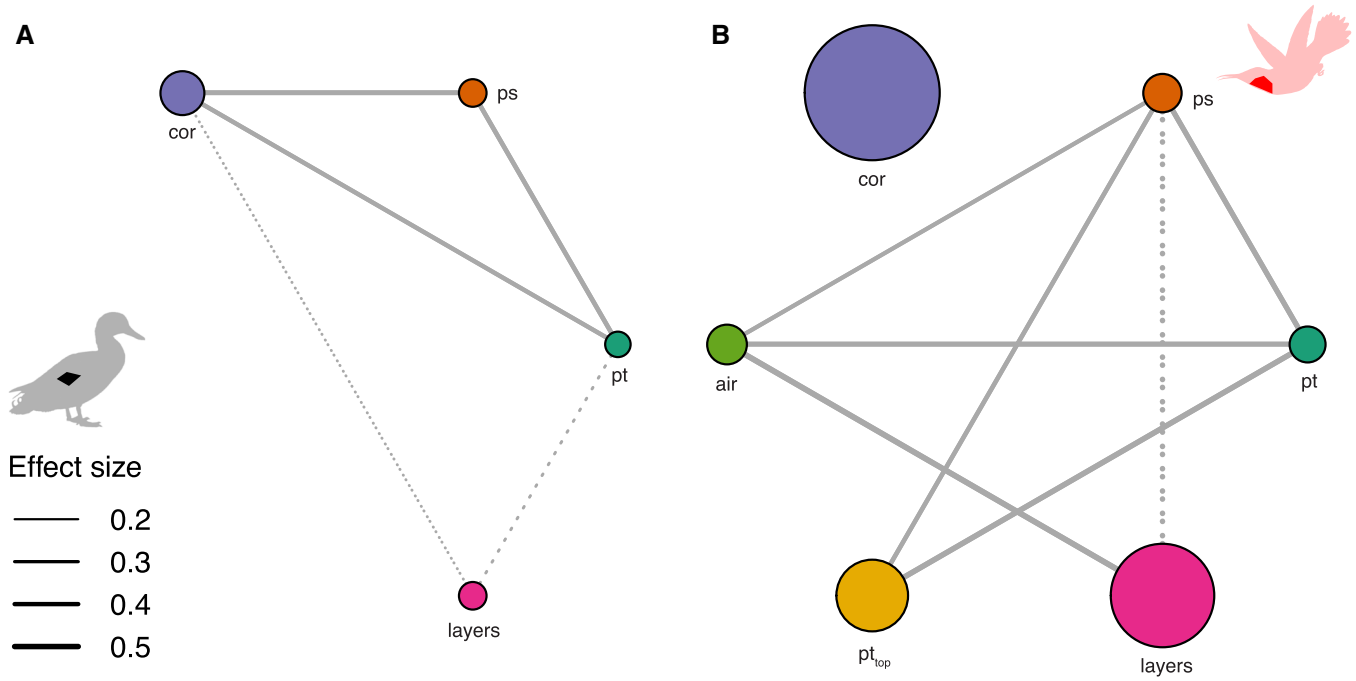

Figure 3. Evolutionary relationships in complex feather nanostructures. Network visualization of evolutionary covariation among nanostructural traits in duck wing feathers (A) and hummingbird gorget feathers (B). Abbreviations correspond to cortex thickness (cor), platelet spacing (ps), platelet thickness (pt), number of platelet layers (layers), top platelet thickness (pt_{top}), and air space diameter (air). Solid lines indicate significantly positive correlations and dashed lines indicate negative relationships among traits. Line thickness indicates strength (effect size) of evolutionary covariance (i.e. absolute value of correlation coefficient; see legend). Only significant connections are shown. Vertex size corresponds to evolutionary disparity (based on an OU model fit) relative to other traits (see Fig. S10 for further details).

Results were similar for crown feathers, with a few notable exceptions. Specifically, standard errors of evolutionary disparity estimates were much larger (expected given the lower sample size, $N = 11$brk; Fig. S10); the cortex and number of layers had significantly lower variances than in gorget feathers (Fig. S10); and spectra were significantly more variable in green wavelengths (Fig. S13B).

These results were generally robust to sampling error. For nanostructural traits, the best-fitting Ornstein Uhlenbeck (OU) model of trait evolution was preferred in 32 and 65 out of 100 jackknife simulations (gorget and crown feathers, respectively). For spectral traits, the preferred OU model was selected in all cases for both feather types. Relative disparities among nanostructural (Fig. S11) and spectral traits (Fig. S13A) were remarkably similar to the full model for gorget feathers. Spectral disparities for crown feathers showed considerably more noise (Figs. S11B, S13B), likely owing to a smaller sample size ($N = 11$) for this body region.

Discussion

Here, we elucidate mechanisms and evolutionary patterns of nanoscale morphology and dynamic colors in charismatic hummingbirds (*Trochilidae*). We identify consistent form-function relationships using optical models, demonstrate that these relationships impact the evolution of both nanostructure and color, and show how morphological complexity facilitates greater evolutionary disparity relative to a clade with simpler nanostructures.

We first show that stacks of melanosomes bounded by keratin in hummingbird feather barbules produce color through multilayer interference, as predicted by Greenewalt et al. (1960a). The novel aspect here is that we consider melanin absorption (Stavenga et al. 2015) and use a parametric model generated from direct TEM measurements to accurately predict reflectance spectrum shape and reflectance peak for ~75% of species (Fig. S6). Interestingly, an optical model not accounting for keratin space between adjacent platelets slightly outperformed the full optical model (Fig. S8), indicating that the amount of keratin between platelets may be less optically important than the thickness of the melanin platelets themselves. Given biological variation, even between and within barbules, and that we only sampled one barbule per species, this degree of agreement is excellent. Discrepancies might arise from sampling of nonrepresentative barbules (i.e., those whose color does not match the overall color of the feather) or from a tilted cross-sectional slice of the barbule during the processing for TEM.

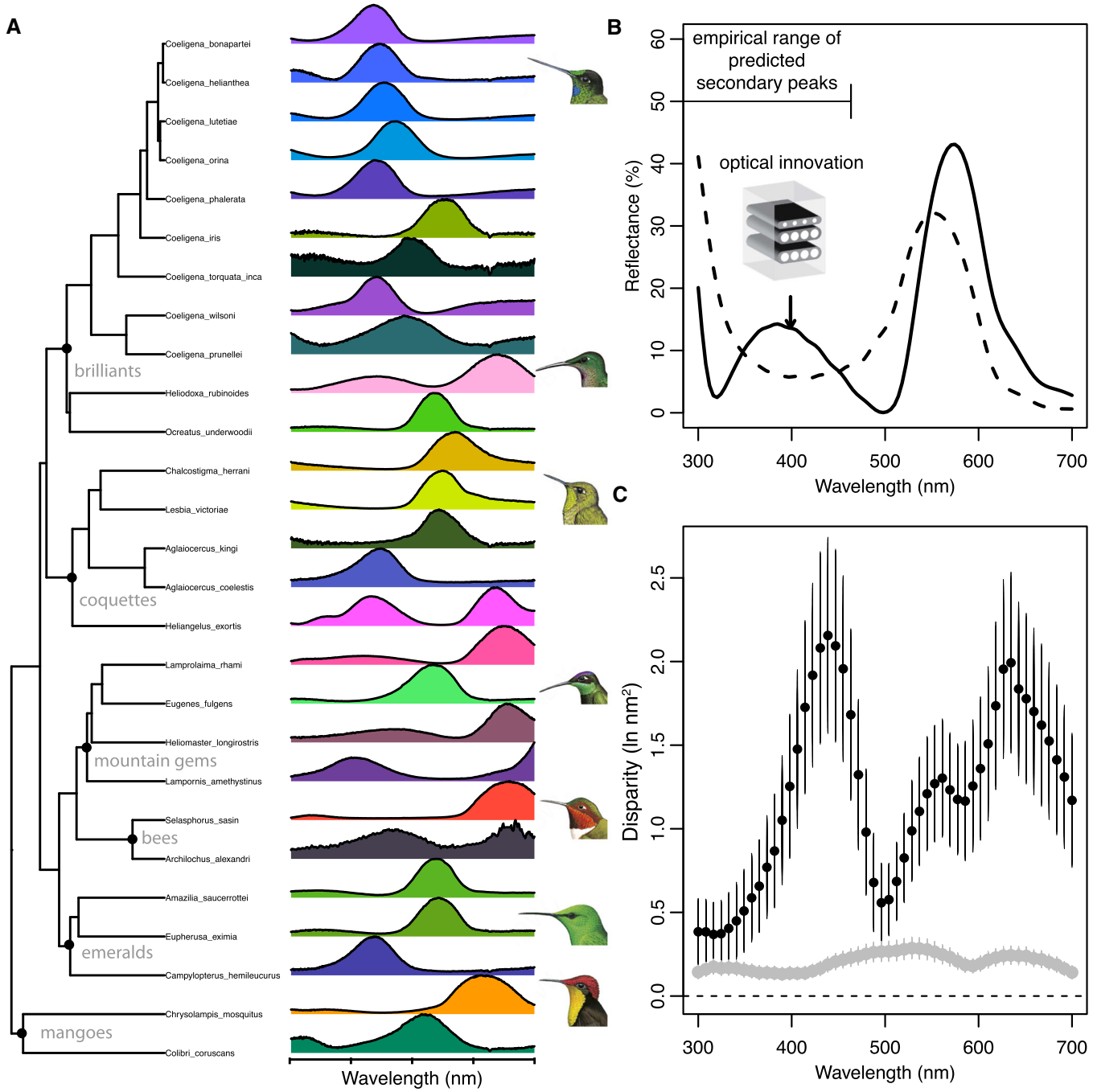

Figure 4. **Evolutionary variation in spectral shape across hummingbirds. (A) Normalized reflectance spectra ($R_{max} = 1$) of gorget feathers ($N = 27$), colored according to human visual models, on a phylogeny of hummingbirds (McGuire et al. 2014). (B) Small, superficial platelets interact with the cortex to cause secondary peaks (arrow). Optical simulations based on the morphology of *Ocreatus underwoodii* (Fig. 1B) with (solid) and without small platelets (dashed line). (C) Evolutionary disparity of iridescent coloration in hummingbird gorget feathers (black) and dabbling duck wing patches (gray). Points show evolutionary variances (±SE) for 50 wavelength bins from ln-transformed reflectance data. Results for hummingbirds based on 27 species illustrated in panel A. Image credits: Handbook of the Birds of the World (Schuchmann 1999).**

Our use of multiple species enabled us to identify variables important to color variation between species. Melanosome platelet thickness strongly impacts hue, as expected under multilayer interference theory in which thicker layers produce longer wavelength colors (Kinoshita 2008). Saturation, or color purity, may decline with air cavity size (Fig. 2B,D) as conditions for an ideal multilayer are approached (Kinoshita 2008). Brightness was not predicted by any nanostructural parameter, but may instead be

influenced by microstructural (rather than nanostructural) traits such as the density, angle, or size of feather barbules (Shawkey et al. 2005). Secondary peaks in some species are likely produced by the optical interaction between the upper cortex and small melanin platelets that lay directly below it (Fig. 4B). The cortex, thus, contributes to both primary and secondary peak production by adding an additional refractive layer.

Evolutionary increases in the number of free morphological parameters comprising a complex trait are thought to increase the potential for morphological diversification (Vermeij 1973). A recent study in dabbling ducks (Anatidae: *Anas*), a monophyletic clade of birds with an isolated iridescent plumage patch (in the wings) and a highly conserved color-producing morphology, showed that form-function relationships explain evolutionary variation in color (Eliason and Shawkey 2012; Eliason et al. 2015). While this work hinted at a role for morphological complexity in opening up opportunity for trait diversification, comparisons among clades differing in morphological complexity are needed to rigorously test Vermeij (1973)'s "biological versatility" hypothesis. Some aspects of hummingbird nanostructures evolve in a correlated manner (Fig. 3B), suggestive either of functional constraints (Maia et al. 2012) or strong correlated selection (Roff and Fairbairn 2012). For example, the negative evolutionary correlation between the number of platelet layers and platelet spacing (Fig. 3B) may be because fewer large platelets can fit in a barbule of a given size. Alternatively, if certain combinations of color attributes (e.g., hue, brightness) are preferred in social or mating contexts, then, over time, this correlated selection could drive genetic correlations between morphological subtraits involved in producing these colors. By contrast, the cortex may be decoupled because its dimensions are independent of other morphological traits, or because its development occurs via a distinct mechanism from the rest of the barbule. Evolutionary decoupling of cortex thickness in hummingbirds (Fig. 3B) but not ducks (Fig. 3A) might stem from differences in how these nanostructures develop. In any case, this evolutionary decoupling of the cortex in hummingbirds allows an additional degree of freedom that we demonstrate has strong optical effects, through its interaction with small platelets at the surface of barbules (Figs. 4B, S7), to drive the explosive color diversification in the clade (Fig. 4A). This is the first study showing the role of the cortex in affecting the color of multilayers from a comparative standpoint. Taken together, our results show that, on average, morphological subtraits in hummingbird nanostructures are more numerous (complex) and independent, resulting in greater overall morphological disparity than those in ducks.

We have here identified proximate mechanisms for color production and variation, and demonstrated a clear link between morphological complexity and evolutionary variation in both morphology and color. How iridescent signals evolve likely depends on both extrinsic social (e.g., female preferences) and ecological factors (e.g., light environment) as well as the intrinsic nature of the structures that produce the colors, in particular, their complexity and modularity. Understanding the proximate basis of these traits will be critical for understanding their ultimate pathways (Eliason 2018).

AUTHOR CONTRIBUTIONS

M.D.S, J.P., and R.M. collected data. M.D.S., R.M., C.M.E. designed the study and analyzed the data; all authors wrote the manuscript.

ACKNOWLEDGMENTS

We thank Carla Cicero (MVZ) for assistance with feather sampling. Reena Zalpuri provided help with TEM and Liliana D'Alba and the rest of the EON lab group, and, John Bates provided helpful feedback on the manuscript. Research funding was provided by a AFOSR Grant FA9550-18-1-0477 and HFSP grants RGY 0083 and RGP 0047 (to M.D.S.), a Bass Postdoctoral Fellowship (to C.M.E.), and a Colombian Administrative Department for Science and Technology – Colciencias Grant code 111571250482—contract number 248–2016 (to J.L.P.).

DATA ARCHIVING

Datasets used in analyses are available on Dryad https://doi.org/10.5061/dryad.jsxksn05h.

LITERATURE CITED

Adams, D. C. 2013. Comparing evolutionary rates for different phenotypic traits on a phylogeny using likelihood. Syst. Biol. 62:181–192.

Adams, D. C., and E. Otárola-Castillo. 2013. Geomorph: an r package for the collection and analysis of geometric morphometric shape data. Methods Ecol. Evol. 4:393–399.

Alfaro, M. E., D. I. Bolnick, and P. C. Wainwright. 2005. Evolutionary consequences of many-to-one mapping of jaw morphology to mechanics in labrid fishes. Am. Nat. 165:E140–E154.

Bartoszek, K., J. Pienaar, P. Mostad, S. Andersson, and T. F. Hansen. 2012. A phylogenetic comparative method for studying multivariate adaptation. J. Theor. Biol. 314:204–215.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67:1–48.

Bleiweiss, R. 1992. Reversed plumage ontogeny in a female hummingbird: implications for the evolution of iridescent colours and sexual dichromatism. Biol. J. Linn. Soc. 47:183–195.

Clark, C. J., and T. J. Feo. 2010. Why do Calypte hummingbirds "Sing" with both their tail and their syrinx? An apparent example of sexual sensory bias. Am. Nat. 175:27–37.

Clark, C. J., J. A. McGuire, E. Bonaccorso, J. S. Berv, and R. O. Prum. 2018. Complex coevolution of wing, tail, and vocal sounds of courting male bee hummingbirds. Evolution 72:630–646.

Clavel, J., L. Aristide, and H. Morlon. 2018. A penalized likelihood framework for high-dimensional phylogenetic comparative methods and an application to new-world monkeys brain evolution. Syst. Biol. 68(1):93–116. https://doi.org/10.1093/sysbio/syy045.

Claverie, T., and S. N. Patek. 2013. Modularity and rates of evolutionary change in a power-amplified prey capture system. Evolution 67:3191–3207.

Denton, J. S. S., and D. C. Adams. 2015. A new phylogenetic test for comparing multiple high-dimensional evolutionary rates suggests interplay

of evolutionary rates and modularity in lanternfishes (Myctophiformes; Myctophidae). Evolution 69:2425–2440.

Diamant, R., A. Garcí-Valenzuela, and M. Fernández-Guasti. 2012. Reflectivity of a disordered monolayer estimated by graded refractive index and scattering models. J. Opt. Soc. Am. A 29:1912–1921.

Dumont, E. R., K. Samadevam, I. Grosse, O. M. Warsi, B. Baird, and L. M. Davalos. 2014. Selection for mechanical advantage underlies multiple cranial optima in new world leaf-nosed bats. Evolution 68:1436–1449.

Durrer, H. 1977. Schillerfarben der vogelfeder als evolutionsproblem. Denkschr. Schweiz. Naturf. Ges. 91:1–127.

Eliason, C. M. 2018. How do complex animal signals evolve? PLoS Biol. 16:e3000093.

Eliason, C. M., and M. D. Shawkey. 2012. A photonic heterostructure produces diverse iridescent colours in duck wing patches. J. R. Soc. Interface 9:2279–2289.

Eliason, C. M., P.-P. Bitton, and M. D. Shawkey. 2013. How hollow melanosomes affect iridescent colour production in birds. Proc. Royal Soc. London Ser.B-Biol. Sci 280:20131505.

Eliason, C. M., R. Maia, and M. D. Shawkey. 2015. Modular color evolution facilitated by a complex nanostructure in birds. Evolution 69:357–367.

Felice, R. N., M. Randau, and A. Goswami. 2018. A fly in a tube: macroevolutionary expectations for integrated phenotypes. Evolution 72:2580–2594.

Garahan, A., L. Pilon, J. Yin, and I. Saxena. 2007. Effective optical properties of absorbing nanoporous and nanocomposite thin films. J. Appl. Phys. 101:014320.

Gingerich, P. D. 2009. Rates of evolution. Annu. Rev. Ecol. Evol. Syst. 40:657–675.

Giraldo, M. A., J. L. Parra, and D. G. Stavenga. 2018. Iridescent colouration of male Anna's hummingbird (*Calypte anna*) caused by multilayered barbules. J. Comp. Physiol. A 204:965–975.https://doi.org/10.1007/s00359-018-1295-8.

Greenewalt, C., W. Brandt, and D. Friel. 1960a. Iridescent colors of hummingbird feathers. J. Opt. Soc. Am. 50:1005–1013.

———. 1960b. The iridescent colors of hummingbird feathers. Proc. Am. Philos. Soc. 104:249–253.

Hadfield, J. D., and S. Nakagawa. 2010. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. J. Evol. Biol. 23:494–508.

Ho, L., and C. Ané. 2014. A linear-time algorithm for gaussian and non-gaussian trait evolution models. Syst. Biol. 63:397–408.

Hoekstra, H. E., J. M. Hoekstra, D. Berrigan, S. N. Vignieri, A. Hoang, C. E. Hill, P. Beerli, and J. G. Kingsolver. 2001. Strength and tempo of directional selection in the wild. Proc. Natl. Acad. Sci. 98:9157–9160.

Hu, D., J. A. Clarke, C. M. Eliason, R. Qiu, Q. Li, M. D. Shawkey, C. Zhao, L. D'Alba, J. Jiang, and X. Xu. 2018. A bony-crested Jurassic dinosaur with evidence of iridescent plumage highlights complexity in early paravian evolution. Nat. Commun. 9:217.

Huang, H., and D. L. Rabosky. 2014. Sexual selection and diversification: reexamining the correlation between dichromatism and speciation rate in birds. Am. Nat. 184:E101–E114.

Hughes, M., S. Gerber, and M. A. Wills. 2013. Clades reach highest morphological disparity early in their evolution. Proc. Natl. Acad. Sci. 110:13875–13879.

Jellison, G. 1993. Data analysis for spectroscopic ellipsometry. Thin Solid Films 234:416–422.

Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, and C. O. Webb. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464.

Kinoshita, S. 2008. Structural colors in the realm of nature. World Scientific, Singapore.

Leal, M., and J. B. Losos. 2015. A naturalist's insight into the evolution of signal redundancy. Am. Nat. 186:ii–iv.

Leertouwer, H. L., B. D. Wilts, and D. G. Stavenga. 2011. Refractive index and dispersion of butterfly chitin and bird keratin measured by polarizing interference microscopy. Opt. Express 19:24061–24066.

Maia, R., R. H. F. Macedo, and M. D. Shawkey. 2012. Nanostructural self-assembly of iridescent feather barbules through depletion attraction of melanosomes during keratinization. J. R. Soc. Interface 9:734–743.

Maia, R., D. R. Rubenstein, and M. D. Shawkey. 2013. Key ornamental innovations facilitate diversification in an avian radiation. Proc. Natl. Acad. Sci. 110:10687–10692.

Mayr, E. 1960. The emergence of evolutionary novelties. Pp. 349–380 *in* Evolution after Darwin. University of Chicago Press, Chicago, IL.

McGuire, J. A., C. C. Witt, J. Remsen J. V., A. Corl, D. L. Rabosky, D. L. Altshuler, and R. Dudley. 2014. Molecular phylogenetics and the diversification of hummingbirds. Curr. Biol. 24:910–916.

Meadows, M. G., N. I. Morehouse, R. L. Rutowski, J. M. Douglas, and K. J. Mcgraw. 2011. Quantifying iridescent coloration in animals: a method for improving repeatability. Behav. Ecol. Sociobiol. 65:1317–1327.

Ord, T. J., D. C. Collar, and T. J. Sanger. 2013. The biomechanical basis of evolutionary change in a territorial display. Funct. Ecol. 27:1186–1200.

Parra, J. L. 2010. Color evolution in the hummingbird genus coeligena. Evolution 64:324–335.

Price, T. D. 2008. Speciation in birds. Roberts & Co., Greenwood Village, CO.

Prum, R. O. 2010. The Lande-Kirkpatrick mechanism is the null model of evolution by intersexual selection: implications for meaning, honesty, and design in intersexual signals. Evolution 64:3085–3100.

Roff, D. A., and D. J. Fairbairn. 2012. A test of the hypothesis that correlational selection generates genetic correlations. Evolution 66:2953–2960.

Rohlf, F. J., and M. Corti. 2000. Use of two-block partial least-squares to study covariation in shape. Syst. Biol. 49:740–753.

Schuchmann, K. 1999. Family trochilidae (Hummingbirds). Pp. 468–680 *in* Handbook of the birds of the world. Lynx Edicons, Spain.

Seddon, N., C. A. Botero, J. A. Tobias, P. O. Dunn, H. E. A. MacGregor, D. R. Rubenstein, J. A. C. Uy, J. T. Weir, L. A. Whittingham, and R. J. Safran. 2013. Sexual selection accelerates signal evolution during speciation in birds. Proc. Royal Soc. London Ser.B-Biol. Sci. 280:20131065.

Servedio, M. R., and R. Burger. 2014. The counterintuitive role of sexual selection in species maintenance and speciation. Proc. Natl. Acad. Sci. 111:8113–8118.

Shawkey, M. D., and L. D'Alba. 2017. Interactions between colour-producing mechanisms and their effects on the integumentary colour palette. Philos. T. R. Soc. B 372:20160536.

Shawkey, M. D., A. M. Estes, L. M. Siefferman, and G. E. hill. 2005. The anatomical basis of sexual dichromatism in non-iridescent ultraviolet-blue structural coloration of feathers. Biol. J. Linn. Soc. 84:259–271.

———. 2003. Nanostructure predicts intraspecific variation in ultraviolet-blue plumage colour. Proc. Royal Soc. London Ser.B-Biol. Sci 270:1455–1460.

Starck, J. M., and R. E. Ricklefs. 1998. Variation, constraint, and phylogeny: comparative analysis of variation in growth. *in* Avian growth and development: evolution within the Altricial-precocial spectrum. Oxford University Press, Oxford, U.K.

Stavenga, D. G., H. L. Leertouwer, D. C. Osorio, and B. D. Wilts. 2015. High refractive index of melanin in shiny occipital feathers of a bird of paradise. Light Sci. Appl. 4:e243.

Stayton, C. T. 2006. Testing hypotheses of convergence with multivariate data: morphological and functional convergence among herbivorous lizards. Evolution 60:824–841.

Stoddard, M. C., and R. O. Prum. 2011. How colorful are birds? Evolution of the avian plumage color gamut. Behav. Ecol. 22:1042–1052.

Van Valen, L. 1974. Multivariate structural statistics in natural history. J. Theor. Biol. 45:235–247.

Vermeij, G. J. 1973. Biological versatility and earth history. Proc. Natl. Acad. Sci. 70:1936–1938.

Wainwright, P. C. 2007. Functional versus morphological diversity in macroevolution. Annu. Rev. Ecol. Evol. Syst. 38:381–401.

West-Eberhard, M. J. 1983. Sexual selection, social competition, and speciation. Q. Rev. Biol. 58:155–183.

Yoshioka, S., S. Kinoshita, H. Iida, and T. Hariyama. 2012. Phase-adjusting layers in the multilayer reflector of a jewel beetle. J. Phys. Soc. Jpn. 81:054801.

Associate Editor: Rebecca Fuller
Handling Editor: Maria Servedio

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Model selection for evolutionary rate models.
Figure S1. Species sampling in hummingbirds (Trochilidae).
Figure S2. Variation in the size of surficial melanin platelets in iridescent hummingbird feather barbules.
Figure S3. Partitioning of variance in morphological traits.
Figure S4. Relationship between hue and angle in hummingbird feathers.
Figure S5. Effect of resolution on reflectance simulations.
Figure S6. Reflectance simulations for the 6-parameter optical model.
Figure S7. Reflectance simulations for different optical models.
Figure S8. Match between simulated and measured reflectance in hummingbirds.
Figure S9. Effects of morphological trait variation on simulated reflectance spectra.
Figure S10. Rates of morphological evolution in hummingbird feather nanostructures.
Figure S11. Species sampling effects on rates of nanostructural evolution.
Figure S12. Species sampling effects on estimates of evolutionary covariation among nanostructural traits.
Figure S13. Species sampling effects on rates of spectral evolution.
Figure S14. Species sampling effects on estimates of spectral covariation.