

# charisma: An R package to perform reproducible color characterization of digital images for biological studies

Shawn T. Schwartz\*†<sup>1,2,3</sup>, Whitney L.E. Tsai\*†<sup>1,4,5</sup>, Elizabeth A. Karan<sup>1</sup>, Mark S. Juhn<sup>1</sup>, Allison J. Shultz<sup>5</sup>, John E. McCormack<sup>4</sup>, Thomas B. Smith<sup>1</sup>, and Michael E. Alfaro<sup>1</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA

<sup>2</sup>Department of Psychology, Stanford University, CA, USA

<sup>3</sup>Wu Tsai Neurosciences Institute, Stanford University, CA, USA

<sup>4</sup>Moore Laboratory of Zoology, Occidental College, CA, USA

<sup>5</sup>Ornithology Department, Natural History Museum of Los Angeles County, CA, USA

\*These authors contributed equally to this work.

†Correspondence concerning this article should be directed to Shawn Schwartz  
([stschwartz@stanford.edu](mailto:stschwartz@stanford.edu)) and/or Whitney Tsai ([wtsai@nhm.org](mailto:wtsai@nhm.org)).

## 1 Acknowledgments

2 We extend our gratitude to current and former members of the Alfaro Lab, Center for Tropical  
3 Research, and Moore Laboratory of Zoology whose feedback greatly influenced the development of  
4 **charisma**. Specifically, we thank Mackenzie Perillo and Kevin Wang for their help in preparing  
5 the testing and validating datasets for **charisma**, Gregory Grether for helpful comments on earlier  
6 versions of this manuscript, and Hannah Weller for insightful and inspiring conversations about  
7 color pattern evolution, computational methods for classifying color, and the development of both  
8 the **colordistance** and **recolorize** R packages, of which **charisma** was inspired by and builds  
9 upon.

## 10 Funding

11 S.T.S. and W.L.E.T. were supported by the University of California, Los Angeles (UCLA), Depart-  
12 ment of Ecology and Evolutionary Biology (EEB). W.L.E.T. was additionally supported by the  
13 UCLA, Dissertation Year Fellowship; UCLA, EEB, Lida Scott Brown Fellowship; and the National  
14 Science Foundation, Graduate Research Fellowship [DGE-2034835]. Any opinions, findings, and  
15 conclusions or recommendations expressed in this material are those of the authors and do not  
16 necessarily reflect the views of the National Science Foundation.

## 17 Authors' Contributions

- 18 • **S.T.S.:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology,  
19 Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing  
20 - Original Draft Preparation, Writing - Review & Editing
- 21 • **W.L.E.T.:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology,  
22 Project Administration, Resources, Supervision, Validation, Visualization, Writing - Original  
23 Draft Preparation, Writing - Review & Editing
- 24 • **E.A.K.:** Conceptualization, Methodology, Writing - Review & Editing
- 25 • **M.S.J.:** Conceptualization, Methodology, Writing - Review & Editing
- 26 • **A.J.S.:** Methodology, Resources, Writing - Review & Editing
- 27 • **J.E.M.:** Resources, Writing - Review & Editing
- 28 • **T.B.S.:** Methodology, Writing - Review & Editing
- 29 • **M.E.A.:** Conceptualization, Methodology, Writing - Review & Editing

30 **Data/Code Availability** 30

- 31 • Analytic data/code for this manuscript can be found on the following Open Science Frame- 31
- 32 work repo: <https://osf.io/cqg59>. 32
- 33 • `charisma` R package on CRAN: <https://cran.r-project.org/package=charisma>. 33
- 34 • `charisma` package GitHub repo: <https://github.com/shawntz/charisma>. 34
- 35 • `charisma` package docs and interactive tutorials: <https://shawnschwartz.com/charisma>. 35

36 **Competing Interests** 36

37 We declare that there are no conflicts of interest with respect to the authorship or publication of 37

38 this article. 38

39 **ORCiD** 39

- 40 • Shawn T. Schwartz  <https://orcid.org/0000-0001-6444-8451> 40
- 41 • Whitney L.E. Tsai  <https://orcid.org/0000-0003-3274-2829> 41
- 42 • Elizabeth A. Karan  <https://orcid.org/0000-0003-1942-0307> 42
- 43 • Mark S. Juhn  <https://orcid.org/0000-0002-3319-1443> 43
- 44 • Allison J. Shultz  <https://orcid.org/0000-0002-2089-4086> 44
- 45 • John E. McCormack  <https://orcid.org/0000-0002-0912-1461> 45
- 46 • Thomas B. Smith  <https://orcid.org/0000-0002-5978-6912> 46
- 47 • Michael E. Alfaro  <https://orcid.org/0000-0002-8898-8230> 47

48 **Author Note.** Portions of this manuscript have appeared in a talk presented by S.T.S. at the 48

49 Society for Integrative and Comparative Biology Annual (Virtual) Meeting on January 3, 2021 49

50 ([Schwartz et al., 2021](#)), as well as in the published dissertation of W.L.E.T. ([Nakashima, 2024](#)). 50

## Abstract

1. Advances in digital imaging and software tools have provided increasingly accessible datasets and methods for analyzing color evolution. Despite the variety of computational packages available, most rely on color classification before running analyses. Previous methods limit the ability to analyze large-scale image databases, decreases the accuracy of downstream analyses, and is not always representative of biologically relevant color classes.

2. Here, we present **charisma**, an R package designed to characterize the distribution of distinct color classes in images suitable for large-scale studies of biological organisms. Here, we demonstrate the utility of our package through an analysis of color evolution in a sample of diverse and charismatic birds, tanagers, in the subfamily Thraupinae.

3. We show that **charisma** can quickly and accurately classify every pixel in an image and validate these results using pre-identified, canonical color swatches. We find that **charisma** color classifications are consistent with those made by color-pattern experts in the field. Applying **charisma** to tanager color evolution, we find that **charisma** outputs seamlessly integrate with downstream evolutionary analyses.

4. Our results demonstrate that using **charisma** to manually curate and characterize colors in images provides a standardized, reliable, and reproducible framework for high-throughput color classification.

**Keywords:** automation · color · color classification · color pattern · macroevolution · R · reproducibility · software package

Anonymized Data/Code for Peer Review

- Analytic data/code for this manuscript can be found on the following Open Science Framework repo:  
<https://osf.io/cqg59/>
  - charisma package GitHub repo: <https://anonymous.4open.science/r/charisma-C87E/>

## 81 1 Introduction

82 For well over a century, biologists have documented and been intrigued by the charismatic coloration  
83 and patterning of earths organisms (Darwin, 1981; Mayr et al., 1963; Santangeli et al., 2023).  
84 Animal colors and patterns have been shown to have ecologically important functions such as intra-  
85 and inter-specific communication in the form of sexual or social signaling, which includes crypsis,  
86 advertisement, or mimicry (Cooney et al., 2019; Feller et al., 2017; Håstad et al., 2005; Irwin, 1994;  
87 Rabosky et al., 2016). The evolution of conspicuous coloration and patterning has historically been  
88 studied using genetic and observational approaches (Andersson and Amundsen, 1997; Barlow et al.,  
89 2018; Ehrlich et al., 1977; McMillan et al., 1999; Neudecker, 1989; Stoddard et al., 2020). However,  
90 innovations in digital imaging and novel software tools have provided increasingly accessible datasets  
91 and reproducible methods for quantitative ecological and evolutionary investigations of color and  
92 pattern (Chan et al., 2019; Endler, 2012; Endler et al., 2018; Hemingson et al., 2024; Maia et al.,  
93 2013; Valvo et al., 2020; Van Belleghem et al., 2018; van den Berg et al., 2020; Weller and Westneat,  
94 2019). To date, these approaches have advanced the field's understanding of the evolution of sexual  
95 dichromatism (i.e., males and females differing in color) and the factors influencing why some  
96 species have more cryptic, conspicuous, or diverse colors (Cooney et al., 2022; Dale et al., 2015;  
97 Shultz and Burns, 2013, 2017; Yu et al., 2024).

98 Despite this progress, eco-evolutionary explanations for why species display specific colors over  
99 others (e.g., blue vs. green) have been lacking. The ability to characterize discrete color classes  
100 in organisms is important to understand the drivers of color evolution and specifically address  
101 questions like: Does habitat influence the colors birds evolve? Which colors provide crypsis or  
102 conspicuousness? Do specific colors play a role in sexual or social signaling? Given the multi-  
103 dimensional, continuous nature of color and color space, it is both conceptually and technically  
104 challenging to carve up the color spectrum into discrete bins, limiting the number of studies (and  
105 software utilities) that use color categorization (Delhey et al., 2023; Ibáñez-Álamo et al., 2025).  
106 Additionally, many color analysis packages require users to input the number of color classes ( $k$ ) *a*  
107 *priori* to compute various downstream color metrics, which means images must first undergo color  
108 profiling such that  $k$  meaningfully captures the specific dominant color classes of each organism  
109 (Maia et al., 2019; Van Belleghem et al., 2018; Weller and Westneat, 2019). Previous analyses have  
110 been limited to small clades with low color variation across species (Alfaro et al., 2019; Hemingson  
111 et al., 2019; Weller and Westneat, 2019). A large-scale color class analysis in birds used a bespoke  
112 method for color categorization to analyze bird illustrations (Delhey et al., 2023). While this allowed  
113 for broad analyses of color evolution in birds, illustrators understandably exhibit artistic license in  
114 their renderings. Color analysis based on images of organisms can reduce these inherent biases.  
115 These small datasets and custom methods limit reproducibility and the breadth of organisms that  
116 can be studied using this approach.

117 To fill this gap, we introduce **charisma**, an R package designed to automatically determine the  
118 presence or absence of key human-visible color categories in images that can be used in studies of  
119 organisms across the tree of life. We provide a flexible and reproducible framework to efficiently  
120 characterize discrete color classes in images. The **charisma** package is suitable for large-scale stud-  
121 ies of color and color pattern evolution and can be seamlessly integrated into popular downstream  
122 quantitative analysis workflows in R (e.g., **geiger**, **patternize**, **pavo**, etc.). We first describe the  
123 workflows for the **charisma** pipeline and then demonstrate the efficacy of our software by (*i*) vali-  
124 dating **charisma**'s color-classification performance using images comprised of predetermined color

125 classes, (*ii*) applying **charisma** to a small set of standardized, museum specimen images of *Tangara* 125  
126 species and allies in the subfamily Thraupinae and comparing the color classification performance 126  
127 of **charisma** to those made by an expert, and (*iii*) demonstrating how color classification labels 127  
128 produced by **charisma** can be leveraged within a trait-based macroevolutionary analysis of color 128  
129 evolution. 129

## 130 2 Materials and Methods 130

### 131 2.1 The **charisma** pipeline 131

132 The primary function of **charisma** is to characterize the distribution of human-visible color classes 132  
133 that are present in an image (**Figure 1**). The **charisma** package was intentionally designed with 133  
134 efficiency and reproducibility in mind, facilitating a standardized and extensible workflow for fully- 134  
135 and semi-automated color classification. To accurately profile the distribution of colors present 135  
136 in an image, we created a biologically inspired color look-up-table (CLUT), which contains hue, 136  
137 saturation, and value (HSV) boundaries for each of 10 human-visible color classes (black, blue, 137  
138 brown, green, grey, orange, purple, red, white, and yellow). Segmentation of continuous color space 138  
139 into useful categories has previously been done and provides a means to investigate colors and color 139  
140 types in evolutionary biology [Delhey et al. \(2023\)](#); [He et al. \(2022\)](#); [Ibáñez-Álamo et al. \(2025\)](#). 140  
141 We determined HSV color space as the best option for the CLUT for a variety of reasons. (*i*) HSV 141  
142 color space provides an intuitive space to separate human-visible color categories. (*ii*) There is a 142  
143 simple calculation to convert between HSV and red, green, blue (RGB) color spaces, which is useful 143  
144 because most digital images are stored in RGB color space and color distance and complexity are 144  
145 easily calculated in RGB color space. In practice, this conversion can be performed directly using 145  
146 R's built-in `rgb2hsv()` function, which returns hue, saturation, and value in normalized units that 146  
147 can then be scaled (e.g., hue  $\times$  360°, saturation  $\times$  100%, value  $\times$  100%) for downstream analyses 147  
148 in **charisma**. (*iii*) An alternative color space is CIELab which was defined by the International 148  
149 Commission on Illumination in 1976 and provides a perceptually uniform color space representative 149  
150 of how the human eye perceives color ([Commission Internationale de l'Éclairage \(CIE\), 2004](#)). 150  
151 However, working in CIELab space is biased towards human visual perception, computationally 151  
152 intensive and requires information about lighting conditions, which limits its utility for large scale 152  
153 analyses of images taken in varying light environments or from different sources. 153

154 To determine the boundaries for each color across the HSV color space, we iteratively developed 154  
155 a schema for defining non-overlapping ranges of hue, saturation, and value triplets for each of 155  
156 the 10 color labels. For each color category, there are distinct HSV ranges that do not overlap 156  
157 and can be validated using the command `charisma::validate()`. These non-overlapping bound- 157  
158aries were tuned leveraging standardized, color-calibrated, research quality photos of bird museum 158  
159 specimens. The CLUT shipped with **charisma** out-of-the-box is open-source and editable (see 159  
160 `charisma::CLUT`). We also provide users the flexibility to define and test their own custom CLUTs 160  
161 (e.g., optimizing classification boundaries for different image sets including non-biological applica- 161  
162 tions) for use within **charisma**. 162

163 Before applying the CLUT-based classification in **charisma**, images are pre-processed using the 163  
164 `recolorize2(bins = 4, cutoff = 20)` function from the R package `recolorize` ([Weller et al., 164  
165 2024](#)). The `recolorize2()` function provides a flexible tool for human-visible color classification, 165

166 which results in a spatially ‘smoothed’ version of the input image by removing stray or noisy pixels  
167 that may negatively influence the downstream **charisma** color-label classification (for an example,  
168 see **Figure 1b**). The **recolorize** package primarily accomplishes this by performing a spatial-color  
169 binning procedure using the RGB ‘histogram’ method (Weller et al., 2024; Weller and Westneat,  
170 2019), which bins each image pixel into a fixed number of color bins. This down-sampled representa-  
171 tion of the dominant image colors enables faster color classification by passing each bin’s RGB triplet  
172 (i.e., the average color of all pixels assigned to each color bin) into **charisma::color2label(r,**  
173 **g,** **b)**. The **charisma::color2label(r, g, b)** function tests for the union of the supplied RGB  
174 triplet within the non-overlapping HSV ranges for each color class in the CLUT and returns the  
175 matching color label from the set of 10 color classes. This approach is computationally efficient  
176 because only a fixed number of bins need to be classified (i.e., 4 bins for each of 3 channels;  $4^3 = 64$   
177 cluster centers using the histogram method; cluster centers with a Euclidean distance less than a  
178 default cutoff of 20 are combined). The alternative approach classifies every single pixel within the  
179 image, which could incorrectly classify shadows or image artifacts and becomes a computationally  
180 intensive procedure as the size and resolution of the image scales.

181 The entire **charisma** workflow can be run either autonomously or with minimal-to-maximal user-  
182 intervention to accommodate a wide range of use cases. This is accomplished by manually merging  
183 the color categories, replacing particular colors, and/or using a threshold cutoff to remove colors  
184 with low proportions of pixels that fall below the specified threshold. We develop a custom im-  
185 plementation that draws on the **recolorize** binning procedure with interactive prompts for color  
186 replacement and merging (Weller et al., 2024), empowering users to check the color classification  
187 at each stage of the **charisma** pipeline in an iterative and flexible fashion. The final **charisma**  
188 output (**Figure 1c**) for each image includes the number of colors present ( $k = 1\dots10$ ), a table with  
189 presence and absence data for each color class, a log of all manual interventions performed, and  
190 R objects that can easily be passed through existing evolutionary analysis pipelines, like those in  
191 the R package **pavo** (Maia et al., 2013, 2019) or analyses of evolutionary models and rates (e.g.,  
192 Harmon et al., 2008; Pennell et al., 2014; Revell, 2012, 2024). Overall, **charisma** facilitates a highly  
193 standardized and reproducible pipeline to characterize color class data from images of organisms for  
194 downstream analyses. By carefully designing the package to handle a spectrum of fully automated  
195 to completely manual workflows, users are equipped to adapt the precision of **charisma**-derived  
196 color label classifications to their particular research needs while enjoying automated data and  
197 image processing provenance out-of-the-box.

## 198 2.2 Validating **charisma**’s color classification performance

199 We validated the performance of **charisma** and the accuracy of our CLUT by first testing known  
200 color datasets. We obtained color grids for each of our 10 colors from the images on the Wikipedia  
201 “Shades of [Color]” pages (**Table S1**). These images contain grids of 9-25 color squares represen-  
202 tative of each color, which we used as input in the automated **charisma** pipeline.

## 203 2.3 Imaging bird museum specimens

204 We then illustrate the utility of **charisma** for evolutionary color analyses with images of tanagers in  
205 the subfamily Thraupinae (Family Thraupidae). Tanagers in this subfamily have been well studied

206 in terms of molecular phylogenetics and color evolution (Burns et al., 2014, 2016; Price-Waldman  
207 et al., 2020; Shultz and Burns, 2017). Previous work has found that lineages within Thraupinae  
208 have the highest evolutionary rates of plumage complexity in tanagers (Price-Waldman et al., 2020).  
209 The subfamily Thraupinae also contains the notably colorful genus *Tangara* (and allied genera that  
210 have been split from the traditional *Tangara* genus), an ideal clade for testing **charisma**. We  
211 photographed 32 bird museum specimens at the Natural History Museum of Los Angeles County  
212 (**Table 3**). This testing dataset consists of incomplete sampling within the subfamily and contains  
213 mostly male specimens with a few female specimens that exhibit minimal sexual dichromatism in the  
214 human visible spectrum. Specimens were photographed under consistent conditions using a Nikon  
215 D70s with Novoflex 35mm lens and Natural LightingNaturesSunlite 30-W full spectrum fluorescent  
216 bulbs. Each image included a color standard and RAW image files were white-balanced before  
217 **charisma** processing. The images were segmented manually by annotating the pixel coordinates  
218 to create precise polygonal mask contours directly around the boundary of each bird body (à la  
219 Schwartz and Alfaro, 2021) to remove the background. To focus our analyses on plumage coloration,  
220 we removed the bill, leg, tag, and cotton eye pixels of the specimen images prior to characterizing  
221 color metrics using **charisma**.

## 222 2.4 Bird coloration datasets

223 To classify color in our bird museum specimen images, we ran both the fully automated **charisma**  
224 pipeline and a modified version where we manually reduced pixel color noise (Figure 1b). For  
225 the fully automated datasets, we tested threshold values of 5%, 7.5%, and 10%, where any color  
226 with a proportion of pixels lower than these threshold values would be removed from the color  
227 classification. We implemented these thresholds to automatically remove colors that may have  
228 been misclassified due to image artifacts like shadows or feather overlap. To ground-truth our  
229 **charisma** classifications, we extracted color classifications for our species from a dataset of color  
230 classifications for all members of the Thraupidae (true tanager) family determined by an expert in  
231 tanager coloration (A.J.S.). **charisma** color classification performance was evaluated by comparing  
232 the categorical classifications derived from the automatic and manual color datasets against our  
233 expert dataset using a binary contingency table (Powers, 2020). We used our expert color dataset  
234 as the true colors of the birds and classified each **charisma**-identified color classification as a true  
235 positive (i.e., ‘hit’), false negative (i.e., ‘miss’), false positive (i.e., ‘false alarm’), or true negative  
236 (i.e., ‘correct rejection’).

## 237 2.5 Evolutionary analyses

238 We used our datasets and a previously published tanager phylogeny (Burns et al., 2014) to explore  
239 variation in the rates of color evolution and reconstruct ancestral color states. Additionally, because  
240 the color-producing mechanisms and structures are well known in birds (Hill, 2006; Mason and  
241 Bowie, 2020; Stoddard and Osorio, 2019; Stoddard and Prum, 2011), we also estimated rates and  
242 reconstructed ancestral states for color-producing mechanisms. We transformed our color data  
243 by grouping discrete colors by avian color-producing mechanism. We designated melanin-based  
244 colors as black, brown, and grey, carotenoid-based colors as red, orange, and yellow, and structural  
245 colors (produced by feather nanostructure, including those with pigment overlays) as green, blue,  
246 and purple (Hill, 2006). For purposes of this study, we have oversimplified the color producing

mechanisms, specifically for green (includes carotenoids and nanostructure), blue, and purple (both include melanin and nanostructure). We removed white from the color mechanism analysis because it has two potential mechanisms: a lack of pigment or feather nanostructure. For each color and mechanism, we used the `fitDiscrete()` function in the R package `geiger` to test the fit of two models of evolution: the equal rates (ER) model, which assumes the rates of gains and losses of a color are equal, and the all rates different (ARD) model, which assumes gains and losses of a color occur at different rates. We compared transition rates for every color in our datasets to determine the potential effects of automated, expert, and manual classifications on evolutionary analyses. We used the sample-size corrected Akaike Information Criterion (AICc) to select the best model for each color and mechanism for our manual dataset. We then leveraged the best-fitting model to reconstruct ancestral states for our manual dataset. Using the R package `phytools`, we estimated maximum-likelihood probabilities for each color and mapped them on the tanager phylogeny (Burnham et al., 1998; Burns et al., 2014; Huelsenbeck et al., 2003; Paradis et al., 2004; Revell, 2012).

## 3 Results

### 3.1 Validation of `charisma`'s color classification performance

**Figure 2** demonstrates `charisma`'s ability to characterize the colors in an image. Here, we find that the majority of `charisma` color classifications identify the pre-assigned color class of the input grid. The highest variation in color classifications was present in the red and orange grids with four different colors being called for each. Brown had the second highest proportion of calls for both red and orange grids, which highlights the general difficulty of delineating between the boundaries of these three colors in HSV color space.

### 3.2 Comparison of color datasets

We found that the manual color dataset outperformed automatic color datasets when compared to the expert color dataset (**Table 4** and **Tables S2, S3, S4**). Black was well-classified in all datasets and brown and grey had high false alarm rates in automatic datasets (**Table 4** and **Tables S2, S3, S4**). The bases and tips of bird feathers, especially contour feathers, are often different colors, with the base generally being black, brown, grey, or white and the tip containing more highly pigmented or structural coloration (Price-Waldman et al., 2025; Terrill and Shultz, 2023). In bird museum specimens, overlapping feathers can be misaligned to create grey or brown patches where the base of feathers show (**Figure 1c**). These color artifacts may contribute to the over-representation of brown and grey in the automatic color datasets. For our analyses, these patches represented unwanted image artifacts, however, depending on the goal of the user, the classification of such subtle variation could be a benefit.

Red and orange had high miss rates in the automatic datasets (**Table 4** and **Tables S2, S3, S4**). Red ( $n = 2$ ) and orange ( $n = 1$ ) are rare in our dataset, but when present, they are represented by very small patches. As expected, the automatic thresholding procedure removed colors with a small proportion of pixels, making underrepresented color categories difficult to classify using the

285 automatic pipeline. Together, the small sample size and small patch size of red and orange in these  
286 birds contribute to the high miss rates in the automatic datasets. The classifications derived from  
287 the manual **charisma** pipeline show an almost identical color profile to the expert color dataset  
288 as evidenced by the near-100% hit and correct rejection rate (**Table 4**). As such, signal-to-noise  
289 ratio (SNR) can be further improved by running **charisma** without the thresholding procedure and  
290 manually correcting both underrepresented color patches to reduce misses and color patches sitting  
291 at the boundary of the CLUT color class boundaries to reduce false alarm identifications.

### 292 3.3 Evolutionary analyses

293 To understand how classifying color in different ways (i.e., automatic vs. manual datasets) impacts  
294 evolutionary rates analyses, we compared evolutionary rates estimated for the expert, manual, and  
295 automatic datasets (**Figure 3** and **Table 1**). We found that the automatic datasets elevated the  
296 rate of evolution of grey across all models providing evidence that uncorrected color artifacts in bird  
297 museum specimens, likely due to misaligned feathers, impact downstream evolutionary analyses.  
298 We also found slower evolutionary rates for blue and green in the automatic datasets, which may  
299 be due to the higher miss rates of **charisma**-identified blue and green color classes (**Table 4** and  
300 **Tables S2, S3, S4**). As predicted, we see high congruence in evolutionary rates between our  
301 expert and manual color datasets.

302 Given the similarity between our expert and manual color datasets, we present the results of our  
303 evolutionary analyses using the best-performing **charisma** classification (i.e., the manual color  
304 dataset). We note that these results and interpretations are based on incomplete sampling of a  
305 subfamily of tanagers and they are presented as an example of the utility of **charisma** for these types  
306 of evolutionary analyses. Additionally, we have categorized highly complicated color-producing  
307 mechanisms into two simple mechanisms. Any biological interpretations of the data should be taken  
308 with caution. We tested two models of evolution on discrete color categories and colors grouped by  
309 color-producing mechanism (see **Figure 4**, **Table 1**, and **Table 2**). We excluded purple and the  
310 melanin mechanism from our analyses because purple was not present in any species and melanin  
311 was ubiquitous across all species. When comparing the evolution of the melanin color using the ER  
312 model, brown showed an elevated rate compared to all other color classes, and black showed the  
313 slowest rate (**Figure 3**). Melanin molecules are structurally robust and are often deposited in the  
314 primary wing and tail feathers of birds to increase durability and resist abrasion in these high-use  
315 feathers (Bonser, 1995). This aligns with our finding that melanin coloration is highly conserved  
316 and indicates the need for structural integrity in bird feathers.

317 For color-producing mechanisms, carotenoid coloration showed a higher rate of evolution than  
318 structural coloration (**Figure 4**). Pigment-based color is more widespread across the avian tree of  
319 life than structural color (Hill and McGraw, 2006). However, across our tanager dataset, structural  
320 colors (green and blue), are present in more species than carotenoid colors (red, orange, and yellow),  
321 which may be contributing to the faster rate of evolution for carotenoid coloration. The structural  
322 color, blue, best fit the ARD model of evolution and showed elevated gain and loss rates compared  
323 to all other colors (**Figure 3** and **Table 2**). Where structural color is present, there is evidence that  
324 color diversity accumulates faster than pigment-based color because of the modularity of layering  
325 pigments and structure (Eliason et al., 2015; Maia et al., 2016). These subtle structural changes at  
326 the nanometer level may result in the production of significantly different colors allowing for rapid  
327 evolution and accumulation of color diversity.

328 Lastly, we found the ER model of evolution to be the best-fitting model for every color and color-  
329 producing mechanism except for blue and orange (**Table 1**). Using the best-fitting model for  
330 each color and mechanism, we reconstructed ancestral states and mapped them across the tanager  
331 phylogeny (**Figure 5**). We found that most ancestral nodes had black, brown, yellow, green, and  
332 blue color states, which were also the most common colors in our dataset. Black, white, brown,  
333 and grey have been shown to be the most common colors across the bird tree of life (Delhey et al.,  
334 2023) and differences in our findings are likely due to the uniquely colorful nature of birds in the  
335 subfamily Thraupinae. The ancestral state reconstructions demonstrate that rates of evolution are  
336 largely driven by losses of colors across internal branches of the tree. Orange fit the ARD model  
337 best with a slightly elevated rate of gain of orange than loss (**Figure 3**). However, orange is only  
338 present in one species in our dataset and this result is driven mainly by the uncertainty of the  
339 presence of orange at the root of the phylogeny (see **Figures 3 and 5**).  
339

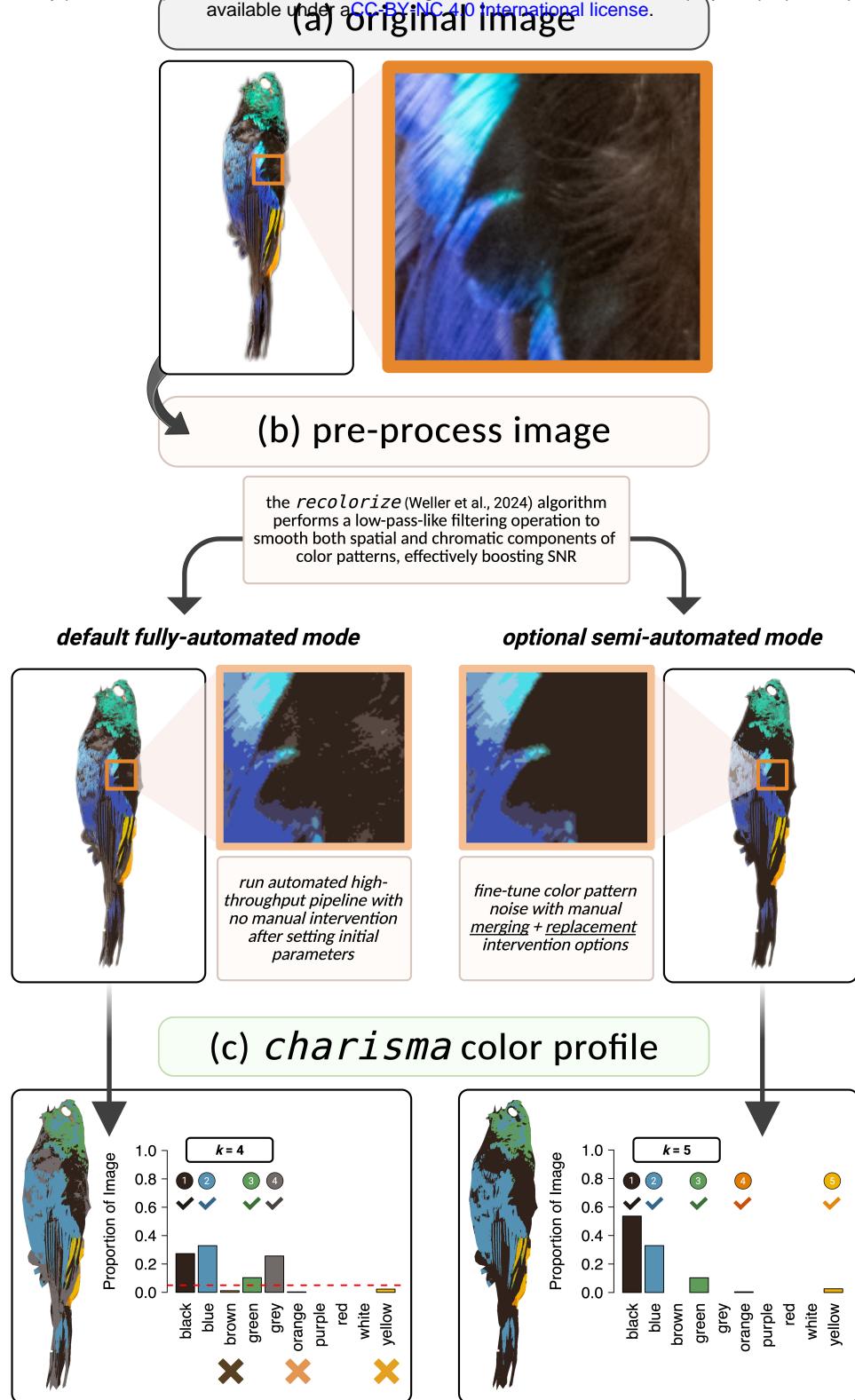
## 340 4 Discussion

341 Here, we present **charisma**, an R package we developed to provide a standardized, reproducible,  
342 and flexible framework for characterizing and classifying color categories from digital images.  
341

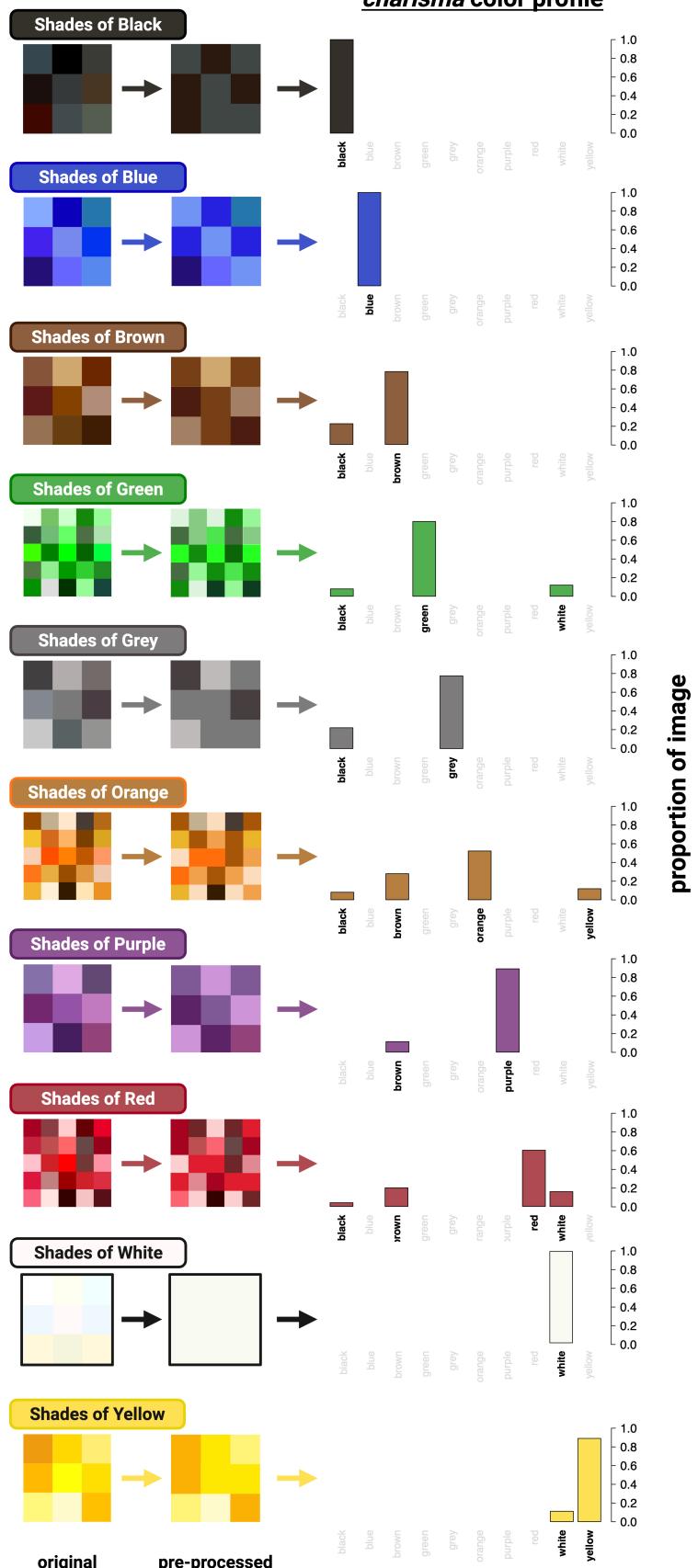
343 To demonstrate its utility, we used color-calibrated images of bird museum specimens and found  
344 that a hybrid approach using both the built-in automated and manual clean-up pipelines provided  
345 the most accurate color profiling. Artifacts from misaligned feathers in our bird museum specimen  
346 image dataset led to over representation of brown and grey in the fully-automatic **charisma**-derived  
347 color classification. While these artifacts might not generalize to all image datasets, we recommend  
348 using **charisma** in two-stages: first, automatically characterize the color profile for each image;  
349 then, manually adjust the outputs using **charisma**'s built-in utilities for classification clean-up to  
350 best fit the biology of your system.  
350

351 Critically, **charisma** is built with a fully open source design philosophy where flexibility and cus-  
352 tomization are both welcomed and encouraged. One example is the built-in functionality to ad-  
353 just the default CLUT based on user needs and their specific image dataset. Cases where this  
354 might be most useful are when fine-tuning is needed for standardized datasets of other organisms,  
355 like J.E. Randall's images of fish (<http://pbs.bishopmuseum.org/images/JER/>), or for extant  
356 datasets comprised of non-standardized images, like those culled from iNaturalist (Matheson, 2014;  
357 <https://www.inaturalist.org>) or eBird (Sullivan et al., 2009).  
357

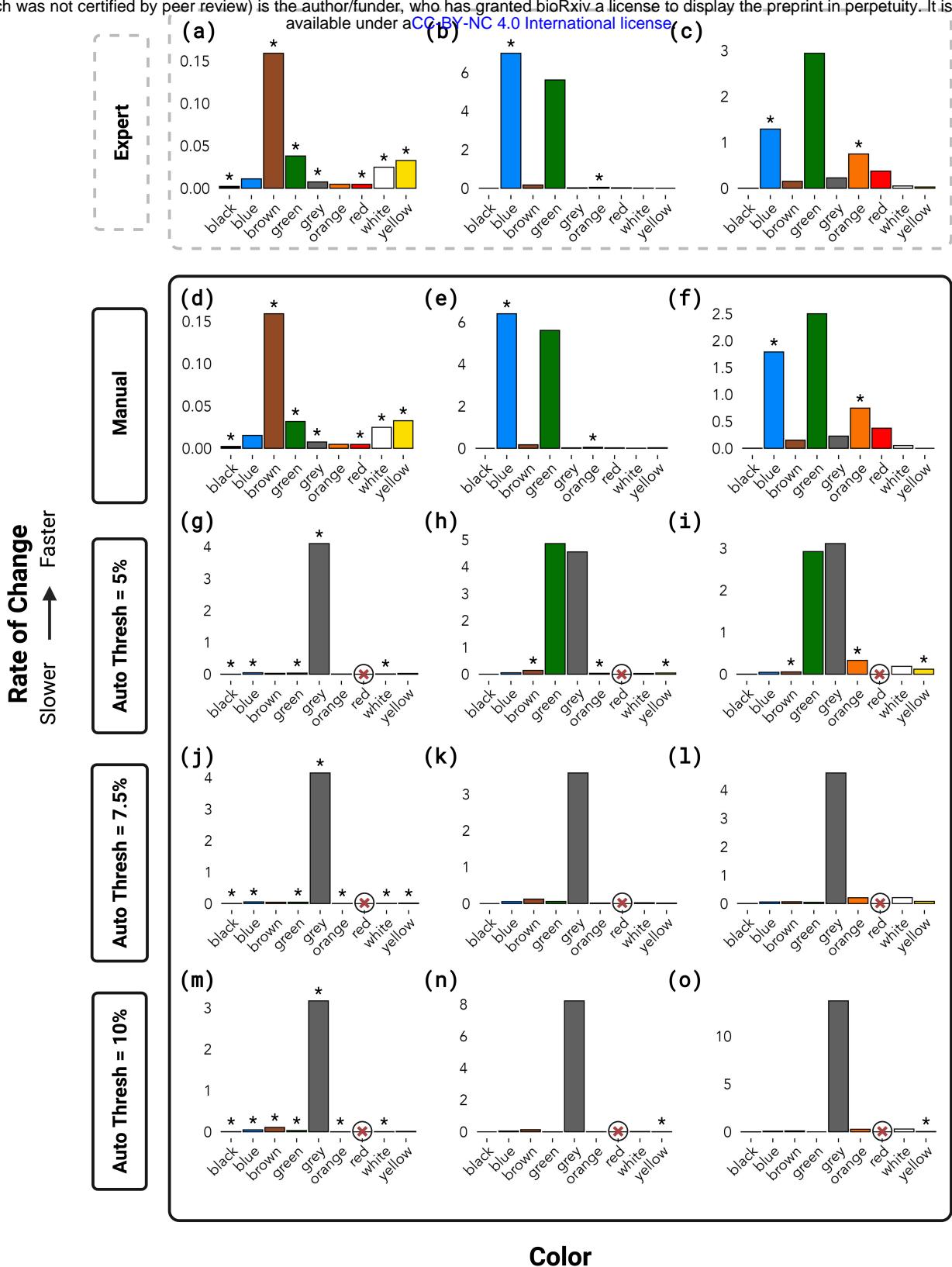
358 Lastly, **charisma** enables seamless transition to downstream evolutionary analyses, which we show-  
359 cased in a vignette on tanagers in the subfamily Thraupinae. More broadly, we present **charisma**  
360 as a solution for high-throughput organismal color analyses with utility that is not restricted to  
361 evolutionary biology applications.  
361



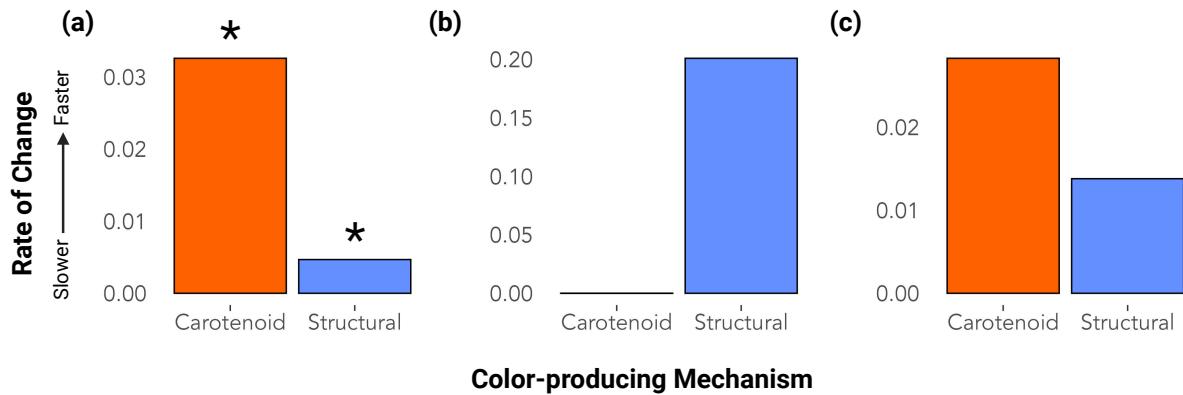
**Fig. 1.** Overview of the *charisma* workflow. **a**, Original digitized specimen image, *Tangara fastuosa*, passed into *charisma* for processing. **b**, The *recolorize* algorithm (see Weller et al., 2024) is used to boost SNR by effectively performing a low-pass-like filter. Then, users are prompted with an option to make manual adjustments to further fine-tune image noise (e.g., biologically irrelevant colors) with interactive color merging and replacement operations (adapted from the *recolorize* R package, see above). **c**, Lastly, users receive the *charisma* “color profile”, which provides a count of the number of discrete color class labels (out of the 10 shown here;  $k$ ) as well as the proportion of pixels in the image that fall into each of the 10 possible color categories. On the left, the dashed red horizontal line indicates a 5% user-specified threshold to perform automated rejection of low frequency color classes. The fully-automated workflow yielded  $k = 4$  (i.e., black, blue, green, and grey). On the right, manual adjustments further reduced noise to eliminate the brown and grey artifacts and correctly call orange and yellow with a 0% threshold. Created in BioRender. (Schwartz and Tsai, 2025a). <https://BioRender.com/bccv8xm>



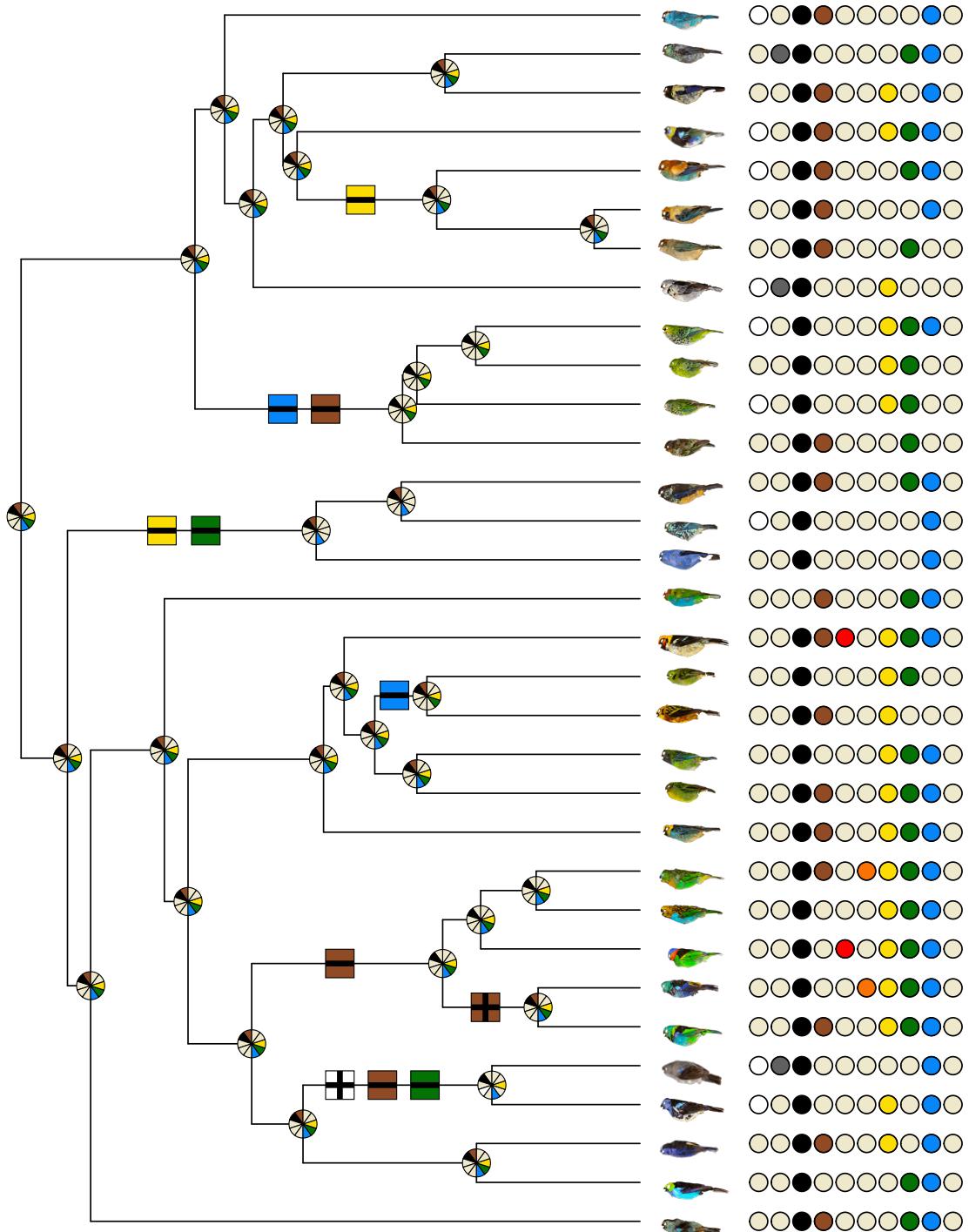
**Fig. 2.** Validating *charisma* color category classifications. Swatches each containing shades of one of nine color categories (white not shown) were obtained from Wikipedia (see **Table S1**) and were then submitted directly to the *charisma* fully-automated pipeline. *charisma* performed well-above chance at identifying each of the nine color categories (see *charisma color profile*). Created in BioRender. (Schwartz and Tsai, 2025b). <https://BioRender.com/ds5tav2>



**Fig. 3.** Evolutionary rates of change for each color and evolutionary model as classified in our expert dataset (A.J.S.) – top-dashed – versus **charisma**-based classifications – **"Manual"**: hybrid classification using a threshold = 0% with manual adjustments (W.L.E.T.); the three remaining classifications were derived from the fully-automated **charisma** pipeline with no manual editing, using increasingly conservative color-proportion thresholds from top-to-bottom. Red  $\otimes$  indicates the color was not called by the **charisma** analysis; asterisks (\*) indicate best fit model for each color category (rates and model selection can be found in **Table 1**). Each plot shows the relative rate of evolutionary change proportional to the highest rate on each plot. Each column represents the following rates: **left**, Equal Rates (ER), **middle**, All Rates Different (ARD) gain, **right**, ARD loss. Note that for the automated dataset with threshold 7.5% (j, k, l), the ER and ARD model selection was inconclusive with an AICc weight of 0.5 for each model. Created in BioRender. ([Schwartz and Tsai, 2025c](#)). <https://BioRender.com/w6sx1dx>



**Fig. 4.** Evolutionary rates for each color-producing mechanism and evolutionary model for the manual dataset: **a**, Equal Rates (ER) rate, **b**, All Rates Different (ARD) gain rate, **c**, ARD loss rate. Asterisks (\*) indicate best fit model for each color-producing mechanism (rates and model selection can be found in **Table 2**). Created in BioRender. (Schwartz and Tsai, 2025d). <https://BioRender.com/8uliuti>



Gain or loss of color at internal branches

Color wheel at each node with ancestral color states  
(≥50% maximum likelihood probability) filled in

**Fig. 5.** charisma color classifications mapped on the tanager phylogeny. For each species, colored dots represent the colors present in the image. Color wheels at each node indicate estimated colors with greater than or equal to 50% maximum likelihood probability. Internal branches are marked with gains or losses of colors. Species order (top to bottom): *Chalcothraupis ruficervix*, *Stilpnia heinei*, *S. cyanoptera*, *S. larvata*, *S. preciosa*, *S. cayana*, *S. vitriolina*, *Poecilotreptus palmeri*, *Ixothraupis guttata*, *I. xanthogastra*, *I. punctata*, *I. rufigula*, *Tangara dowii*, *T. nigroviridis*, *T. vassorii*, *T. gyrola*, *T. parzudakii*, *T. florida*, *T. arthus*, *T. johannae*, *T. schrankii*, *T. xanthocephala*, *T. desmaresti*, *T. cyanoventris*, *T. cyanocephala*, *T. fastuosa*, *T. seledon*, *T. ornata*, *T. mexicana*, *T. velia*, *T. chilensis*, *T. labradorides*.

**Table 1.** Rates of color evolution and ancestral state reconstruction model selection for all datasets.

	Color	Equal Rates Model			All Rates Different Model			
		Rate of gain and loss	AICc	AICc weight	Rate of gain	Rate of loss	AICc	AICc weight
<b>Expert</b>	Black	0.00226*	10.16	0.755	0	0.00226	12.41	0.245
	White	0.0248*	38.98	0.554	0.00859	0.0505	39.42	0.446
	Grey	0.00749*	23.75	0.565	0.0241	0.227	24.27	0.435
	Brown	0.159*	46.21	0.75	0.169	0.152	48.41	0.25
	Red	0.00466*	18.48	0.608	0.0252	0.374	19.36	0.392
	Orange	0.00475	21.4	0.266	0.0499*	0.747*	19.37	0.734
	Yellow	0.0327*	41.28	0.616	0	0.0283	42.23	0.384
	Green	0.0380*	44.87	0.589	5.61	2.94	45.6	0.411
	Blue	0.0111	32.8	0.42	6.99*	1.29*	32.15	0.58
<b>Manual</b>	Black	0.00226*	10.16	0.755	0	0.0026	12.41	0.245
	White	0.0248*	38.98	0.554	0.00859	0.0505	39.42	0.446
	Grey	0.00749*	23.75	0.565	0.0241	0.227	24.27	0.435
	Brown	0.159*	46.21	0.75	0.169	0.152	48.41	0.25
	Red	0.00466*	18.48	0.608	0.0252	0.374	19.36	0.392
	Orange	0.00475	21.4	0.266	0.0499*	0.747*	19.37	0.734
	Yellow	0.0326*	41.28	0.616	0.028	0	42.23	0.384
	Green	0.0317*	44.06	0.514	5.6	2.5	44.16	0.486
	Blue	0.0152	37.3	0.591	6.39*	1.79*	38.03	0.409
<b>Auto Thresh - 5%</b>	Black	0.00226*	10.16	0.755	0	0.0026	12.41	0.245
	White	0.01*	27.84	0.569	0.0278	0.188	28.4	0.431
	Grey	4.1*	46.49	0.64	4.55	3.11	47.64	0.36
	Brown	0.0282	42.59	0.409	0.144*	0.0563*	41.85	0.591
	Red	NA	NA	NA	NA	NA	NA	NA
	Orange	0.00743	24.55	0.467	0.0345*	0.33*	24.29	0.533
	Yellow	0.0243	41.845	0.486	0.0462*	0.12*	41.74	0.514
	Green	0.0359*	44.87	0.72	4.86	2.92	46.75	0.28
	Blue	0.0472*	44.92	0.747	0.055	0.046	47.09	0.253
<b>Auto Thresh - 7.5%</b>	Black	0.00464*	15.63	0.749	0	0.00463	17.82	0.251
	White	0.00747*	23.61	0.579	0.0223	0.209	24.25	0.421
	Grey	4.15*	46.49	0.709	3.58	4.6	48.27	0.291
	Brown	0.0434	45	0.5	0.123	0.0648	45	0.5
	Red	NA	NA	NA	NA	NA	NA	NA
	Orange	0.00478*	18.34	0.619	0.0141	0.205	19.31	0.381
	Yellow	0.0152*	34.74	0.534	0.0131	0.0769	36.02	0.466
	Green	0.0441*	45.36	0.741	0.0582	0.0453	47.46	0.259
	Blue	0.0544*	45.25	0.757	0.0537	0.0567	47.51	0.243
<b>Auto Thresh - 10%</b>	Black	0.00464*	15.63	0.749	0	0.00429	17.82	0.251
	White	0.00482*	18.94	0.551	0.0198	0.292	19.35	0.449
	Grey	3.17*	46.49	0.532	8.21	13.69	46.75	0.468
	Brown	0.107*	45.8	0.659	0.136	0.0948	47.12	0.341
	Red	NA	NA	NA	NA	NA	NA	NA
	Orange	0.00229*	11.83	0.675	0.00876	0.267	13.29	0.325
	Yellow	0.0111	29.32	0.475	0.00748*	0.0347*	29.12	0.525
	Green	0.0317*	40.99	0.63	0.00318	0	42.05	0.37
	Blue	0.0486*	44.85	0.706	0.0484	0.0687	46.61	0.294

Note. Asterisks (\*) indicate best fit model for each color category.

**Table 2.** Rates of color-producing mechanism evolution and ancestral state reconstruction model selection for the manual dataset.

	Equal Rates Model				All Rates Different Model			
	Mechanism	Rate of gain and loss	AICc	AICc weight	Rate of gain	Rate of loss	AICc	AICc weight
Carotenoid		0.0326*	41.28	0.616	0	0.0283	42.23	0.384
Structural		0.00468*	17.35	0.725	0.201	0.0138	19.29	0.275

Note. Asterisks (\*) indicate best fit model for each color category.

**Table 3.** Bird museum specimens.

Catalog Number	Genus	Species	Common Name
LACM 34880	<i>Chalcothraupis</i>	<i>ruficervix</i>	Golden-naped Tanager
LACM 4843	<i>Ixothraupis</i>	<i>guttata</i>	Speckled Tanager
LACM 43685	<i>Ixothraupis</i>	<i>punctata</i>	Spotted Tanager
LACM 16611	<i>Ixothraupis</i>	<i>rufigula</i>	Rufous-throated Tanager
LACM 50758	<i>Ixothraupis</i>	<i>xanthogastra</i>	Yellow-bellied Tanager
LACM 37481	<i>Poecilotreptus</i>	<i>palmeri</i>	Gray-and-gold Tanager
LACM 39216	<i>Stilpnia</i>	<i>cayana</i>	Burnished-buff Tanager
LACM 37716	<i>Stilpnia</i>	<i>cyanoptera</i>	Black-headed Tanager
LACM 29385	<i>Stilpnia</i>	<i>heinei</i>	Black-capped Tanager
LACM 4841	<i>Stilpnia</i>	<i>larvata</i>	Golden-headed Tanager
LACM 53462	<i>Stilpnia</i>	<i>preciosa</i>	Chestnut-backed Tanager
LACM 36850	<i>Stilpnia</i>	<i>vitriolina</i>	Scrub Tanager
LACM 59610	<i>Tangara</i>	<i>arthus</i>	Golden Tanager
LACM 59272	<i>Tangara</i>	<i>chilensis</i>	Paradise Tanager
LACM 59652	<i>Tangara</i>	<i>cyancephala</i>	Red-necked Tanager
LACM 27866	<i>Tangara</i>	<i>cyanoventris</i>	Gilt-edged Tanager
LACM 28775	<i>Tangara</i>	<i>desmaresti</i>	Brassy-breasted Tanager
LACM 16290	<i>Tangara</i>	<i>dowii</i>	Spangle-cheeked Tanager
LACM 60421	<i>Tangara</i>	<i>fastuosa</i>	Seven-colored Tanager
LACM 60422	<i>Tangara</i>	<i>florida</i>	Emerald Tanager
LACM 43661	<i>Tangara</i>	<i>gyrola</i>	Bay-headed Tanager
LACM 30414	<i>Tangara</i>	<i>inornata</i>	Plain-colored Tanager
LACM 37463	<i>Tangara</i>	<i>johannae</i>	Blue-whiskered Tanager
LACM 59219	<i>Tangara</i>	<i>labradorides</i>	Metallic-green Tanager
LACM 32721	<i>Tangara</i>	<i>mexicana</i>	Turquoise Tanager
LACM 40998	<i>Tangara</i>	<i>nigroviridis</i>	Beryl-spangled Tanager
LACM 29400	<i>Tangara</i>	<i>parzudakii</i>	Flame-faced Tanager
LACM 50757	<i>Tangara</i>	<i>schrankii</i>	Green-and-gold Tanager
LACM 53515	<i>Tangara</i>	<i>seledon</i>	Green-headed Tanager
LACM 29453	<i>Tangara</i>	<i>vassori</i>	Blue-and-black Tanager
LACM 43655	<i>Tangara</i>	<i>velia</i>	Opal-rumped Tanager
LACM 33255	<i>Tangara</i>	<i>xanthocephala</i>	Saffron-crowned Tanager

**Table 4.** Performance comparison between the *automated charisma* classification workflow (with manual adjustments) and expert classification across color categories. Each column reports the total number of hits (true positives), misses (false negatives), false alarms (false positives), and correct rejections (true negatives) in the *charisma* dataset. The percentage of correct hits (parentheses) are calculated as the number of hits divided by the total number of color targets (hits + misses) in the expert dataset. The percentage of correct rejections (parentheses) are calculated as the number of correct rejections divided by the total number of non-targets (correct rejections + false alarms) in the expert dataset.

Color	Hits	Misses	Correct Rejections	False Alarms
Black	31 (100%)	0	1 (100%)	0
Blue	25 (93%)	2	5 (100%)	0
Brown	17 (100%)	0	15 (100%)	0
Green	21 (100%)	0	10 (91%)	1
Grey	3 (100%)	0	29 (100%)	0
Orange	2 (100%)	0	30 (100%)	0
Purple	0 (100%)	0	32 (100%)	0
Red	2 (100%)	0	30 (100%)	0
White	9 (100%)	0	23 (100%)	0
Yellow	19 (100%)	0	13 (100%)	0

## 362 References

362

- 363 Alfaro, M. E., Karan, E. A., Schwartz, S. T., and Shultz, A. J. (2019). The evolution of color 363  
364 pattern in butterflyfishes (chaetodontidae). *Integrative and comparative biology*, 59(3):604–615. 364
- 365 Andersson, S. and Amundsen, T. (1997). Ultraviolet colour vision and ornamentation in 365  
366 bluethroats. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 366  
367 264(1388):1587–1591. 367
- 368 Barlow, A., Cahill, J. A., Hartmann, S., Theunert, C., Xenikoudakis, G., Fortes, G. G., Paijmans, 368  
369 J. L., Rabeder, G., Frischauf, C., Grandal-d'Anglade, A., et al. (2018). Partial genomic survival 369  
370 of cave bears in living brown bears. *Nature ecology & evolution*, 2(10):1563–1570. 370
- 371 Bonser, R. H. (1995). Melanin and the abrasion resistance of feathers. *The Condor*, 97(2):590–591. 371
- 372 Burnham, K. P., Anderson, D. R., Burnham, K. P., and Anderson, D. R. (1998). *Practical use of 372  
373 the information-theoretic approach*. Springer. 373
- 374 Burns, K. J., Shultz, A. J., Title, P. O., Mason, N. A., Barker, F. K., Klicka, J., Lanyon, S. M., and 374  
375 Lovette, I. J. (2014). Phylogenetics and diversification of tanagers (passeriformes: Thraupidae), 375  
376 the largest radiation of neotropical songbirds. *Molecular Phylogenetics and Evolution*, 75:41–77. 376
- 377 Burns, K. J., Unitt, P., and Mason, N. A. (2016). A genus-level classification of the family thraup- 377  
378 idae (class aves: Order passeriformes). *Zootaxa*, 4088(3):329–354. 378
- 379 Chan, I. Z., Stevens, M., and Todd, P. A. (2019). Pat-geom: a software package for the analysis of 379  
380 animal patterns. *Methods in Ecology and Evolution*, 10(4):591–600. 380
- 381 Commission Internationale de l'Éclairage (CIE) (2004). Colorimetry. CIE Publication 15:2004, 381  
382 CIE Central Bureau, Vienna, Austria. 382
- 383 Cooney, C. R., He, Y., Varley, Z. K., Nouri, L. O., Moody, C. J., Jardine, M. D., Liker, A., Székely, 383  
384 T., and Thomas, G. H. (2022). Latitudinal gradients in avian colourfulness. *Nature Ecology & 384  
385 Evolution*, 6(5):622–629. 385
- 386 Cooney, C. R., Varley, Z. K., Nouri, L. O., Moody, C. J., Jardine, M. D., and Thomas, G. H. (2019). 386  
387 Sexual selection predicts the rate and direction of colour divergence in a large avian radiation. 387  
388 *Nature communications*, 10(1):1–9. 388
- 389 Dale, J., Dey, C. J., Delhey, K., Kempenaers, B., and Valcu, M. (2015). The effects of life history 389  
390 and sexual selection on male and female plumage colouration. *Nature*, 527(7578):367–370. 390
- 391 Darwin, C. (1981). *The descent of man, and selections in relation to sex*. 391
- 392 Delhey, K., Valcu, M., Muck, C., Dale, J., and Kempenaers, B. (2023). Evolutionary pre- 392  
393 dictors of the specific colors of birds. *Proceedings of the National Academy of Sciences*, 393  
394 120(34):e2217692120. 394
- 395 Ehrlich, P., Talbot, F., Russell, B., and Anderson, G. (1977). The behaviour of chaetodontid fishes 395  
396 with special reference to Lorenz's poster colouration hypothesis. *Journal of Zoology*, 183(2):213– 396  
397 228. 397
- 398 Eliason, C. M., Maia, R., and Shawkey, M. D. (2015). Modular color evolution facilitated by a 398  
399 complex nanostructure in birds. *Evolution*, 69(2):357–367. 399
- 400 Endler, J. A. (2012). A framework for analysing colour pattern geometry: adjacent colours. *Bio- 400  
401 logical Journal of the Linnean Society*, 107(2):233–253. 401

- 402 Endler, J. A., Cole, G. L., and Kranz, A. M. (2018). Boundary strength analysis: Combining 402  
403 colour pattern geometry and coloured patch visual properties for use in predicting behaviour and 403  
404 fitness. *Methods in Ecology and Evolution*, 9(12):2334–2348. 404
- 405 Feller, K. D., Jordan, T. M., Wilby, D., and Roberts, N. W. (2017). Selection of the intrinsic polar- 405  
406 ization properties of animal optical materials creates enhanced structural reflectivity and camou- 406  
407 flage. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1724):20160336. 407
- 408 Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., and Challenger, W. (2008). Geiger: investi- 408  
409 gating evolutionary radiations. *Bioinformatics*, 24(1):129–131. 409
- 410 Håstad, O., Victorsson, J., and Ödeen, A. (2005). Differences in color vision make passerines less 410  
411 conspicuous in the eyes of their predators. *Proceedings of the National Academy of Sciences*, 411  
412 102(18):6391–6394. 412
- 413 He, Y., Varley, Z. K., Nouri, L. O., Moody, C. J., Jardine, M. D., Maddock, S., Thomas, G. H., 413  
414 and Cooney, C. R. (2022). Deep learning image segmentation reveals patterns of uv reflectance 414  
415 evolution in passerine birds. *Nature Communications*, 13(1):5068. 415
- 416 Hemingson, C. R., Cowman, P. F., and Bellwood, D. R. (2024). Analysing biological colour patterns 416  
417 from digital images: An introduction to the current toolbox. *Ecology and Evolution*, 14(3):e11045. 417
- 418 Hemingson, C. R., Cowman, P. F., Hodge, J. R., and Bellwood, D. R. (2019). Colour pattern 418  
419 divergence in reef fish species is rapid and driven by both range overlap and symmetry. *Ecology 419  
420 letters*, 22(1):190–199. 420
- 421 Hill, G. E. (2006). *Bird coloration, volume 2: function and evolution*. Harvard University Press. 421
- 422 Hill, G. E. and McGraw, K. J. (2006). *Bird coloration, volume 1: mechanisms and measurements*, 422  
423 volume 1. Harvard University Press. 423
- 424 Huelsenbeck, J. P., Nielsen, R., and Bollback, J. P. (2003). Stochastic mapping of morphological 424  
425 characters. *Systematic biology*, 52(2):131–158. 425
- 426 Ibáñez-Álamo, J. D., Delhey, K., Izquierdo, L., Valcu, M., and Kempenaers, B. (2025). Colourful 426  
427 urban birds: Bird species successful in urban environments have more elaborate colours and less 427  
428 brown. *Ecology Letters*, 28(4):e70106. 428
- 429 Irwin, R. E. (1994). The evolution of plumage dichromatism in the new world blackbirds: social 429  
430 selection on female brightness. *The American Naturalist*, 144(6):890–907. 430
- 431 Maia, R., Eliason, C. M., Bitton, P.-P., Doucet, S. M., and Shawkey, M. D. (2013). pavo: an r 431  
432 package for the analysis, visualization and organization of spectral data. *Methods in Ecology and 432  
433 Evolution*, 4(10):906–913. 433
- 434 Maia, R., Gruson, H., Endler, J. A., and White, T. E. (2019). pavo 2: new tools for the spectral 434  
435 and spatial analysis of colour in r. *Methods in Ecology and Evolution*, 10(7):1097–1107. 435
- 436 Maia, R., Rubenstein, D. R., and Shawkey, M. D. (2016). Selection, constraint, and the evolution 436  
437 of coloration in african starlings. *Evolution*, 70(5):1064–1079. 437
- 438 Mason, N. A. and Bowie, R. C. (2020). Plumage patterns: Ecological functions, evolutionary 438  
439 origins, and advances in quantification. *The Auk*, 137(4):ukaa060. 439
- 440 Matheson, C. A. (2014). inaturalist. *Reference Reviews*, 28(8):36–38. 440
- 441 Mayr, E. et al. (1963). Animal species and evolution. *Animal species and evolution*. 441

- 442 McMillan, W. O., Weigt, L. A., and Palumbi, S. R. (1999). Color pattern evolution, assortative 442  
443 mating, and genetic differentiation in brightly colored butterflyfishes (chaetodontidae). *Evolution*, 443  
444 53(1):247–260. 444
- 445 Nakashima, W. L. T. (2024). *Genomic approaches to studying evolution and adaptation in birds*. 445  
446 University of California, Los Angeles. 446
- 447 Neudecker, S. (1989). Eye camouflage and false eyespots: chaetodontid responses to predators. In 447  
448 *The butterflyfishes: success on the coral reef*, pages 143–158. Springer. 448
- 449 Paradis, E., Claude, J., and Strimmer, K. (2004). Ape: analyses of phylogenetics and evolution in 449  
450 r language. *Bioinformatics*, 20(2):289–290. 450
- 451 Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., 451  
452 Alfaro, M. E., and Harmon, L. J. (2014). geiger v2. 0: an expanded suite of methods for fitting 452  
453 macroevolutionary models to phylogenetic trees. *Bioinformatics*, 30(15):2216–2218. 453
- 454 Powers, D. M. (2020). Evaluation: from precision, recall and f-measure to roc, informedness, 454  
455 markedness and correlation. *arXiv preprint arXiv:2010.16061*. 455
- 456 Price-Waldman, R. M., Ali, J. R., Shultz, A. J., Hogan, B. G., and Stoddard, M. C. (2025). Hidden 456  
457 white and black feather layers enhance plumage coloration in tanagers and other songbirds. 457  
458 *Science Advances*, 11(30):eadw5857. 458
- 459 Price-Waldman, R. M., Shultz, A. J., and Burns, K. J. (2020). Speciation rates are correlated with 459  
460 changes in plumage color complexity in the largest family of songbirds. *Evolution*, 74(6):1155– 460  
461 1169. 461
- 462 Rabosky, A. R. D., Cox, C. L., Rabosky, D. L., Title, P. O., Holmes, I. A., Feldman, A., and 462  
463 McGuire, J. A. (2016). Coral snakes predict the evolution of mimicry across new world snakes. 463  
464 *Nature communications*, 7(1):1–9. 464
- 465 Revell, L. J. (2012). phytools: an r package for phylogenetic comparative biology (and other things). 465  
466 *Methods in ecology and evolution*, (2):217–223. 466
- 467 Revell, L. J. (2024). phytools 2.0: an updated r ecosystem for phylogenetic comparative methods 467  
468 (and other things). *PeerJ*, 12:e16505. 468
- 469 Santangeli, A., Haukka, A., Morris, W., Arkkila, S., Delhey, K., Kempenaers, B., Valcu, M., Dale, 469  
470 J., Lehikoinen, A., and Mammola, S. (2023). What drives our aesthetic attraction to birds? *njp* 470  
471 *Biodiversity*, 2(1):20. 471
- 472 Schwartz, S. T. and Alfaro, M. E. (2021). Sashimi: A toolkit for facilitating high-throughput organ- 472  
473 ismal image segmentation using deep learning. *Methods in Ecology and Evolution*, 12(12):2341– 473  
474 2354. 474
- 475 Schwartz, S. T., Tsai, W., Karan, E. A., and Alfaro, M. E. (2021). Charisma: An r tool to automati- 475  
476 cally determine discrete color classes for high-throughput color pattern analysis. In *Integrative* 476  
477 *and Comparative Biology*, volume 61, pages E806–E806. Oxford Univ Press Inc Journals Dept, 477  
478 2001 Evans Rd, Cary, Nc 27513 USA. 478
- 479 Schwartz, S. T. and Tsai, W. L. (2025a). Figure 1: Overview of the charisma workflow. Created in 479  
480 BioRender. <https://biorender.com/bccv8xm>. 480
- 481 Schwartz, S. T. and Tsai, W. L. (2025b). Figure 2: Validating charisma color category classifications. 481  
482 Created in BioRender. <https://biorender.com/ds5tav2>. 482
- 483 Schwartz, S. T. and Tsai, W. L. (2025c). Figure 3: Evolutionary rates of change for each color and 483  
484 evolutionary model. Created in BioRender. <https://biorender.com/w6sx1dx>. 484

- 485 Schwartz, S. T. and Tsai, W. L. (2025d). Figure 4: Evolutionary rates for each color-producing 485  
486 mechanism and evolutionary model. Created in BioRender. <https://biorender.com/8uliuti>. 486
- 487 Shultz, A. J. and Burns, K. J. (2013). Plumage evolution in relation to light environment in a novel 487  
488 clade of neotropical tanagers. *Molecular Phylogenetics and Evolution*, 66(1):112–125. 488
- 489 Shultz, A. J. and Burns, K. J. (2017). The role of sexual and natural selection in shaping patterns of 489  
490 sexual dichromatism in the largest family of songbirds (aves: Thraupidae). *Evolution*, 71(4):1061– 490  
491 1074. 491
- 492 Stoddard, M. C., Eyster, H. N., Hogan, B. G., Morris, D. H., Soucy, E. R., and Inouye, D. W. 492  
493 (2020). Wild hummingbirds discriminate nonspectral colors. *Proceedings of the National Academy 493  
494 of Sciences*, 117(26):15112–15122. 494
- 495 Stoddard, M. C. and Osorio, D. (2019). Animal coloration patterns: linking spatial vision to 495  
496 quantitative analysis. *The American Naturalist*, 193(2):164–186. 496
- 497 Stoddard, M. C. and Prum, R. O. (2011). How colorful are birds? evolution of the avian plumage 497  
498 color gamut. *Behavioral Ecology*, 22(5):1042–1052. 498
- 499 Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., and Kelling, S. (2009). ebird: 499  
500 A citizen-based bird observation network in the biological sciences. *Biological conservation*, 500  
501 142(10):2282–2292. 501
- 502 Terrill, R. S. and Shultz, A. J. (2023). Feather function and the evolution of birds. *Biological 502  
503 Reviews*, 98(2):540–566. 503
- 504 Valvo, J. J., Rodd, F. H., Aponte, J. D., Daniel, M., Dwinell, K., Houle, D., and Hughes, K. A. 504  
505 (2020). Colormesh: A novel method for quantifying variation in complex color patterns. *bioRxiv*. 505
- 506 Van Belleghem, S. M., Papa, R., Ortiz-Zuazaga, H., Hendrickx, F., Jiggins, C. D., Owen McMillan, 506  
507 W., and Counterman, B. A. (2018). patternize: An r package for quantifying colour pattern 507  
508 variation. *Methods in Ecology and Evolution*, 9(2):390–398. 508
- 509 van den Berg, C. P., Troscianko, J., Endler, J. A., Marshall, N. J., and Cheney, K. L. (2020). 509  
510 Quantitative colour pattern analysis (qcpa): a comprehensive framework for the analysis of 510  
511 colour patterns in nature. *Methods in Ecology and Evolution*, 11(2):316–332. 511
- 512 Weller, H. I., Hiller, A. E., Lord, N. P., and Van Belleghem, S. M. (2024). recolorize: An r package 512  
513 for flexible colour segmentation of biological images. *Ecology Letters*, 27(2):e14378. 513
- 514 Weller, H. I. and Westneat, M. W. (2019). Quantitative color profiling of digital images with earth 514  
515 movers distance using the r package colordistance. *PeerJ*, 7:e6398. 515
- 516 Yu, J., Duan, H., Zhang, B., Zhang, L., and He, J. (2024). Urbanization alters the geographic 516  
517 patterns of passerine plumage color in china. *Landscape and Urban Planning*, 248:105101. 517