Received: 28 November 2025 | Accepted: 7 April 2026

DOI: 10.1111/2041-210x.70310

Methods in Ecology and Evolution

APPLICATION

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

Shawn T. Schwartz[1,2,3] | Whitney L. E. Tsai[1,4,5] | Elizabeth A. Karan[1] | Mark S. Juhn[1] | Allison J. Shultz[5] | John E. McCormack[4] | Thomas B. Smith[1,6] | Michael E. Alfaro[1]

[1]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, California, USA; [2]Department of Psychology, Stanford University, Stanford, California, USA; [3]Wu Tsai Neurosciences Institute, Stanford University, Stanford, California, USA; [4]Moore Laboratory of Zoology, Occidental College, Los Angeles, California, USA; [5]Ornithology Department, Natural History Museum of Los Angeles County, Los Angeles, California, USA and [6]Institute of Environment and Sustainability, University of California, Los Angeles, Los Angeles, California, USA

Correspondence
Shawn T. Schwartz
Email: stschwartz@stanford.edu

Whitney L. E. Tsai
Email: wtsai@nhm.org

Funding information
University of California, Los Angeles (UCLA); National Science Foundation, Grant/Award Number: DGE-2034835 and DBI-2410072

Handling Editor: Pascal Title

Abstract

1. Advances in digital imaging and software tools have provided increasingly accessible datasets and methods for analysing colour evolution. Despite the variety of computational packages available, most rely on colour classification before running analyses. Previous methods to characterize colour limit the ability to analyse large-scale image databases and are not always representative of biologically relevant colour classes, which decrease the accuracy of downstream analyses.
2. Here, we present `charisma`, an R package designed to characterize the distribution of distinct colour classes in images suitable for large-scale studies of biological organisms. We demonstrate the utility of our package through an analysis of colour 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 colour swatches. We find that `charisma` colour classifications are consistent with those made by colour pattern experts in the field. Applying `charisma` to tanager colour evolution, we find that `charisma` outputs seamlessly integrate with downstream evolutionary analyses.
4. Our results demonstrate that using `charisma` to manually curate and characterize colours in images provides a standardized, reliable and reproducible framework for high-throughput colour classification.

Shawn T. Schwartz and Whitney L. E. Tsai contributed equally to this work.

KEYWORDS
automation, colour, colour classification, colour pattern, macroevolution, R, reproducibility, software package

1 | INTRODUCTION

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

Despite this progress, eco-evolutionary explanations for why species display specific colours over others (e.g. blue vs. green) have been lacking. The ability to characterize discrete human-visible colour classes in organisms is important to understand the drivers of colour evolution and specifically address questions like: Does habitat influence the colours birds evolve? Which colours provide crypsis or conspicuousness? Do specific colours play a role in sexual or social signalling? In birds, colour-producing mechanisms are well known for each colour; thus, knowing the colour classes present in a species can also inform how these mechanisms evolve (Hill & McGraw, 2006; Porzio & Mota, 2025). Given the multidimensional and continuous nature of colour and colour space, it is both conceptually and technically challenging to carve up the colour spectrum into discrete bins, limiting the number of studies and software utilities that use colour classification (Table 1; Delhey et al., 2023; Ibáñez-Álamo et al., 2025; Nicholson et al., 2007; Senior et al., 2022). Additionally, many colour analysis packages require users to input the number of colour classes (k) a priori to compute various downstream colour metrics, which means images must first undergo colour profiling such that `k` meaningfully captures the specific dominant colour classes of each organism (Table 1; Maia et al., 2019; Van Belleghem et al., 2018; Weller & Westneat, 2019). Previous analyses have been limited to small clades with low colour variation across species (Alfaro et al., 2019; Hemingson et al., 2019; Weller & Westneat, 2019). Large-scale colour class analyses in birds have used bespoke methods for colour classification to analyse bird illustrations (Delhey et al., 2023; Senior et al., 2022). While this allows for broad analyses of colour evolution in birds, illustrators exhibit artistic licence in their renderings. Colour analysis based on images of organisms can reduce these inherent biases. These small datasets and custom methods limit the reproducibility and breadth of organisms that can be studied using this approach.

To fill this gap, we introduce `charisma`, an R package designed to streamline the process of characterizing discrete colour classes in digital images that can be used in studies of organisms across the tree of life. We provide a flexible and reproducible framework to efficiently determine the presence or absence of key human-visible colour classes in images. The `charisma` package is suitable for large-scale studies of colour and colour pattern evolution and can be seamlessly integrated into popular downstream quantitative analysis workflows in R (e.g. `geiger`, `patternize`, `pavo`, etc.). We first describe the `charisma` pipeline and then demonstrate the efficacy of our software by (i) validating `charisma`'s colour classification performance using images comprised of predetermined colour classes, (ii) applying `charisma` to a small set of standardized, museum specimen images of *Tangara* species and allies in the subfamily Thraupinae and comparing the colour classification performance of `charisma` to those made by an expert and (iii) demonstrating how colour classification labels produced by `charisma` can be leveraged within a trait-based macroevolutionary analysis of colour evolution.

2 | MATERIALS AND METHODS

2.1 | The `charisma` pipeline

The primary function of `charisma` is to characterize the distribution of human-visible colour classes that are present in an image (Figure 1). The `charisma` package was intentionally designed with efficiency and reproducibility in mind, facilitating a standardized and extensible pipeline for colour classification. Segmentation of continuous colour space into useful classes has previously been done and provides a means to investigate colours and colour types in evolutionary biology (Delhey et al., 2023; Ibáñez-Álamo et al., 2025; Senior et al., 2022). To accurately profile the distribution of colours present in an image, we divided hue, saturation and value (HSV) colour space into each of 10 human-visible colour classes including all recognized primary and secondary colours (red, orange, yellow, green, blue and purple) and black, brown, grey and white (Figure S1). We determined HSV colour space as the best option for the `charisma` package for a variety of reasons. (i) HSV colour space

TABLE 1 Comparison of existing colour classification methods.

Description of colour classification method	Considerations	Source	Benefits of `charisma`
Bespoke: analysis of bird illustrations in CIELab space. A 3D grid of chromatic and achromatic cells was split and merged into 12 colour classes	Bespoke methods are not easily modified or reproduced. Illustrations can be artist biased and may not be available for all organisms	Delhey et al. (2023), Ibáñez-Álamo et al. (2025)	Flexible and reproducible method that can be used on a variety of image datasets with modifications made simple through the CLUT editor and semi-automated mode
Bespoke: analysis of bird illustrations in HSV colour space. Hue was divided into 30 degree bins with exceptions for orange, brown, dark and light for a total of 15 colour classes		Senior et al. (2022)	
Bespoke: analysis of dewlap pattern, colour and size. Individual colour classification from literature and personal images	Bespoke methods are not easily modified or reproduced. Individual colour classifications are not reproducible or feasible for large datasets	Nicholson et al. (2007)	
Histogram: divides colour space into regions and computes the proportion of pixels and average pixel value in each region to produce histograms of unique bin centres based on input image	Single colour classes may be broken up across multiple boundaries	`colordistance`, `recolorize`	The number of potential colour classes is user set and the boundaries can be adjusted based on your biological question and system
K-means clustering: popular method that partitions pixels into a specified number of bins to minimize difference between the pixel cluster centres	Fairly slow and biased towards dominant colours. The number of clusters is usually defined a priori	`colordistance`, `patternize`, `pavo`, `recolorize`	Relatively quick, unbiased towards dominant colours and not a priori *k*-value necessary. The resulting *k*-value can be used as input to these other programs
Receptor noise limited filtering and clustering: sequentially groups of segment adjacent pixels according to specific thresholds of spatial and chromatic acuity	Requires manual designation of thresholds, but can be adjusted. Specific to spatio-chromatic colour pattern analysis	`Quantitative Colour Pattern Analysis`	General workflow that can be used on its own or integrated into existing colour pattern analysis pipelines (e.g. `patternize`, `pavo`)

provides an intuitive space for separating human-visible colour classes and investigating colour class and type evolution regardless of an organism's visual perception. Incorporating an organism's visual perception is beyond the scope of this package and requires spectrophotometric measurements or hyperspectral camera imaging and different analyses (Garcia et al., 2015; Hogan & Stoddard, 2024; van den Berg et al., 2020). (ii) There is a simple calculation to convert between HSV and red, green, blue (RGB) colour spaces, which is useful because most digital images are stored in RGB colour space and colour distance and complexity are easily calculated in RGB colour space. In practice, this conversion can be performed directly using R's built-in `rgb2hsv()` function, which returns hue, saturation and value in normalized units that can then be scaled (e.g. hue × 360°, saturation × 100%, value × 100%) for downstream analyses in `charisma`. (iii) An alternative colour space is CIELab which was defined by the International Commission on Illumination in 1976 and provides a perceptually uniform colour space representative of how the human eye perceives colour (Commission Internationale de l'Éclairage [CIE], 2004). However, working in CIELab space is computationally intensive and requires information about lighting conditions, which limits its utility for large-scale analyses of images taken in varying light environments or from different sources.

We created a biologically inspired colour lookup table (CLUT), which contains distinct, non-overlapping hue, saturation and value boundaries for each colour class (Figure S1). To determine the boundaries for each colour across the HSV colour space, we iteratively developed a schema for defining non-overlapping ranges of hue, saturation and value triplets for each of the 10 colour labels. First, we divided hue into 30 degree bins corresponding to known primary, secondary and tertiary colours (similar to Senior et al., 2022). We merged and adjusted bins to fit our six hue-driven colour classes (red, orange, yellow, green, blue and purple). We defined black, brown, grey and white on the basis of saturation and value. Then, we used `charisma`'s automated mode to adjust and fine-tune the colour boundaries for our standardized, colour-calibrated, research quality photos of bird museum specimens. The CLUT shipped with `charisma` out-of-the-box is open source and editable (see `charisma::CLUT`). It can be validated using the command `charisma::validate()` to ensure that no colour class boundaries are overlapping and that the entire colour spectrum is covered. The interactive CLUT editor app, available on the documentation website (https://charisma.shawnschwartz.com/app), allows for simple optimization of the CLUT by adjusting colour boundaries and adding or subtracting colour classes.

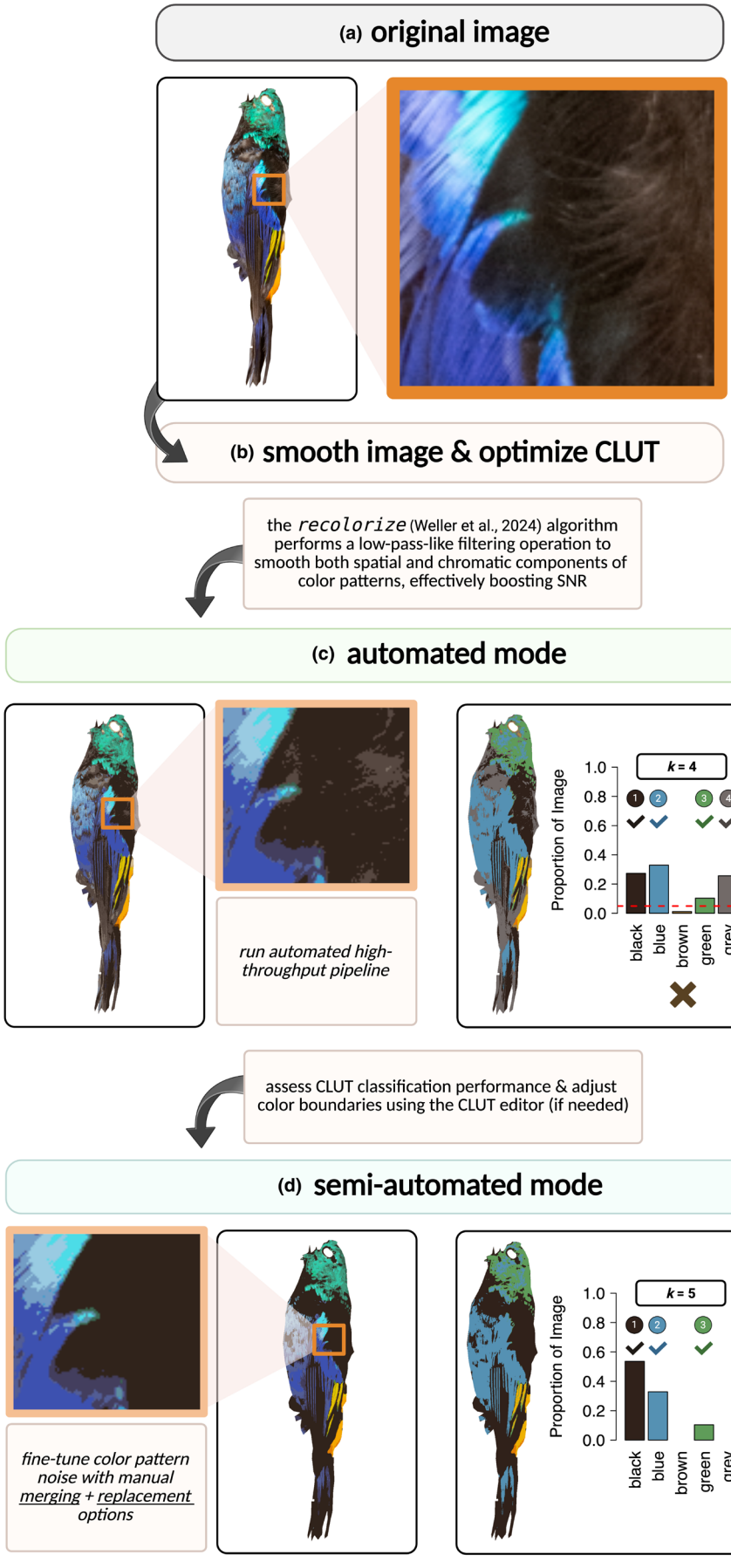

FIGURE 1 Overview of the *charisma* pipeline. (a) Original digitized specimen image, *Tangara fastuosa*, passed into *charisma* for processing. (b) The *recolorize* algorithm (see Weller et al., 2024) is used to boost SNR by effectively performing a low-pass-like filter. (c) Users run the automated mode to check CLUT performance and adjust the colour boundaries according to the image dataset and biological questions using the CLUT editor. (d) Lastly, users are prompted through the semi-automated mode to make optional, manual adjustments to further fine-tune image noise (e.g. biologically irrelevant colours) with interactive colour merging and replacement operations (adapted from the `recolorize` R package, see above). Users receive the *charisma* 'colour profile' (right panel in c and d), which provides a count of the number of discrete colour 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 colour classes. On the left, the dashed red horizontal line indicates a user-specified threshold (5% here) to perform automated rejection of low frequency colour classes. In this example, the automated workflow yielded $k = 4$ (i.e. black, blue, green and grey). The semi-automated mode further reduced noise to eliminate the brown and grey artefacts and correctly call orange and yellow with a 0% threshold.

The CLUT defines colour boundaries such that a given HSV triplet is classified as belonging to a colour category if its hue, saturation and value simultaneously fall within at least one of the colour category's defined range sets. Moreover, colour categories can specify multiple disjunctive range sets (joined by commas), where each set defines a rectangular subregion of HSV space. Within a single range set, the pipe operator (`|`) signifies a union across non-contiguous intervals within a single dimension (e.g. hue ranges of 0–15° or 300–360° for red, which wraps around the hue circle). This syntax enables definition of complex, non-contiguous colour category boundaries. For instance, the category 'brown' perceptually occupies a non-convex region of colour space; therefore, 'brown' in the CLUT is defined as the union of several narrow HSV subregions spanning different value and saturation levels. Our interactive CLUT editor provides users the flexibility to define and test their own custom CLUTs (e.g. optimizing classification boundaries for different image sets including non-biological applications) for use within `charisma`.

Before applying the CLUT-based classification in `charisma`, the `recolorize2()` function from the `R` package `recolorize` (Weller et al., 2024) is used to produce a spatially 'smoothed' version of the input image by removing stray or noisy pixels that may negatively influence the downstream `charisma` colour-label classification (for an example, see Figure 1). The `recolorize` package primarily accomplishes this by performing a spatial-colour binning procedure using the RGB 'histogram' method (Weller et al., 2024; Weller & Westneat, 2019), which bins each image pixel into a fixed number of colour bins. This down-sampled representation of the dominant image colours enables faster colour classification by passing each bin's RGB triplet (i.e. the average colour of all pixels assigned to each colour bin) into `charisma::color2label(r, g, b)`. The `charisma::color2label(r, g, b)` function tests for the union of the supplied RGB triplet within the non-overlapping HSV ranges for each colour class in the CLUT and returns the matching colour label from the set of 10 colour classes. This approach is computationally efficient because only a fixed number of bins need to be classified (i.e. 4 bins for each of 3 channels; $4^3 = 64$ cluster centres using the histogram method; cluster centres with a Euclidean distance less than a default cut-off of 20 are combined). The alternative approach classifies every single pixel within the image, which could incorrectly classify shadows or image artefacts and becomes a computationally intensive procedure as the size and resolution of the image scales.

For best results, we recommend using `charisma` with a multi-stage approach, specifically running both automated and semi-automated workflows together to facilitate accurate colour classification fine-tuned for any given set of specimens (Figure 1; Box 1). The automated mode allows users to quickly run all or a subset of images in their dataset to assess the performance of the default CLUT. If users are not satisfied with the `charisma` outputs, they should iteratively use the CLUT editor app to adjust the colour boundaries and check the results in automated mode. Then, users can use the semi-automated mode to manually merge colours, replace colours and/or use a threshold cut-off to remove colours with low proportions of pixels that fall below the specified threshold. We developed a custom implementation that draws on the `recolorize` binning procedure with interactive prompts for colour replacement and merging (Weller et al., 2024), empowering users to check the colour classification at each stage of the `charisma` pipeline in an iterative and flexible fashion. The final `charisma` output (Figure 1c,d) for each image includes the number of colours present ($k = 1 \ldots 10$), a table with presence and absence data for each colour class, a log of all manual interventions performed and R objects that can easily be passed through existing evolutionary analysis pipelines, like those in the R package `pavo` (Maia et al., 2013, 2019) or analyses of evolutionary models and rates (e.g. Harmon et al., 2008; Pennell et al., 2014; Revell, 2012, 2024). Overall, `charisma` facilitates a highly standardized and reproducible pipeline to characterize colour class data from images of organisms for downstream analyses. By carefully designing the package to handle a spectrum of fully automated to semi-automated workflows, users are equipped to adapt the precision of `charisma`-derived colour label classifications to their particular research needs while enjoying streamlined data and image processing provenance out-of-the-box.

2.2 | Validating `charisma`'s colour classification performance

We validated the performance of `charisma` and the accuracy of our CLUT by first testing known colour datasets. We obtained colour grids for each of our 10 colours from the images on the Wikipedia 'Shades of [Color]' pages (Table S1). These images contain grids of 9–25 colour squares representative of each colour, which we used as input in the automated `charisma` workflow (see Figure 2). We also

BOX 1 Recommended workflow for `charisma` end-users

Step 1: Image dataset generation and pre-processing

(a) We highly recommend generating an image dataset by taking standardized images of organisms with the same camera, lighting conditions and colour standard (like the X-Rite ColorChecker). We understand that there are advantages to using pre-existing images for colour analyses, in which case, we recommend using images from as few sources as possible and only retaining images that can reasonably be compared. For more complete guides on digital imaging for biological studies of colour, see Garcia et al. (2015), Stevens et al. (2007).

(b) White-balance standardized images using a grey standard. For non-standardized image datasets, more complex colour calibration and image standardization is necessary (Akkaynak et al., 2014; Stevens et al., 2007).

(c) Segment images manually or use a pre-existing segmentation toolkit like sashimi (Schwartz & Alfaro, 2021) or PixelMator Pro (https://www.apple.com/pixelmator-pro).

Step 2: Optimizing the `charisma` colour lookup table (CLUT)

(a) Run images through the automated `charisma` workflow to assess the CLUT classification performance. If the image dataset is less than 50 images, run all images. If the image dataset is much larger, run a subset of the images that includes the breadth of colour variation across the dataset.

(b) Assess `charisma` classification outputs and adjust CLUT colour boundaries using the CLUT Editor, if necessary.

(c) Repeat automated `charisma` run and CLUT adjustments until satisfied with CLUT performance.

Step 3: Finalize `charisma` outputs

(a) Run all images through the semi-automated `charisma` workflow and manually adjust images using the merge and replacement options, whenever necessary. When merging and replacing colours, think carefully about your biological system question to ensure accurate representation of colours.

tested images of beetles and fish (Figure S3) using the automated mode to provide examples of images that were not used to tune the CLUT (Alfaro et al., 2019; Karan et al., 2021, 2025; Schwartz & Alfaro, 2021; Weller et al., 2024).

2.3 | Imaging bird museum specimens

We then illustrate the utility of `charisma` for evolutionary colour analyses with images of tanagers in the subfamily Thraupinae (Family Thraupidae). Tanagers in this subfamily have been well studied in terms of molecular phylogenetics and colour evolution (Burns et al., 2014, 2016; Price-Waldman et al., 2020; Shultz & Burns, 2017). Previous work has found that lineages within Thraupinae have the highest evolutionary rates of plumage complexity in tanagers (Price-Waldman et al., 2020). The subfamily Thraupinae also contains the notably colourful genus *Tangara* (and allied genera that have been split from the traditional *Tangara* genus), an ideal clade for testing `charisma`. We photographed 32 bird museum specimens at the Natural History Museum of Los Angeles County (Table 2). This testing dataset consists of incomplete sampling within the subfamily and contains mostly male specimens with a few female specimens that exhibit minimal sexual dichromatism in the human-visible spectrum. Specimens were photographed under consistent conditions using a Nikon D70s with Novoflex 35 mm lens and Natural LightingNaturesSunlite 30-W full spectrum fluorescent bulbs. Each image included an X-Rite ColorChecker standard and we used the 18% neutral grey standard to white-balance the RAW image files in Photoshop before `charisma` processing (Figure S2). The images were manually segmented by annotating the pixel coordinates to create precise polygonal mask contours directly around the boundary of each bird body (à la Schwartz & Alfaro, 2021) to remove the background. To focus our analyses on plumage coloration, we removed the bill, leg, tag and cotton eye pixels of the specimen images prior to characterizing colour metrics using `charisma`.

2.4 | Bird coloration datasets

To classify colour in our bird museum specimen images, we ran both the fully automated `charisma` workflow followed by the semi-automated workflow to manually reduce pixel colour noise (Figure 1). For the fully automated datasets, we tested threshold values of 5%, 7.5% and 10%, where any colour with a proportion of pixels lower than these threshold values would be removed from the colour classification. We implemented these thresholds to automatically remove colours that may have been misclassified due to image artefacts like shadows or feather overlap. For the semi-automated dataset, we merged and/or replaced colours to further refine colour clusters for more accurate downstream colour classification. To ground-truth our `charisma` classifications, we extracted colour classifications for our species from an independent dataset of colour classifications for all members of the Thraupidae (true tanager) family determined by an expert in tanager coloration (A.J.S.). `charisma` colour classification performance was evaluated by comparing the classifications derived from the automated and semi-automated colour datasets against our expert dataset using a binary contingency table (Powers, 2020). We used our expert colour dataset as the true colours of the birds and classified each `charisma`-identified colour classification as a true positive (i.e. 'hit'), false negative (i.e. 'miss'), false positive (i.e. 'false alarm') or true negative (i.e. 'correct rejection').

BRITISH ECOLOGICAL SOCIETY Methods in Ecology and Evolution

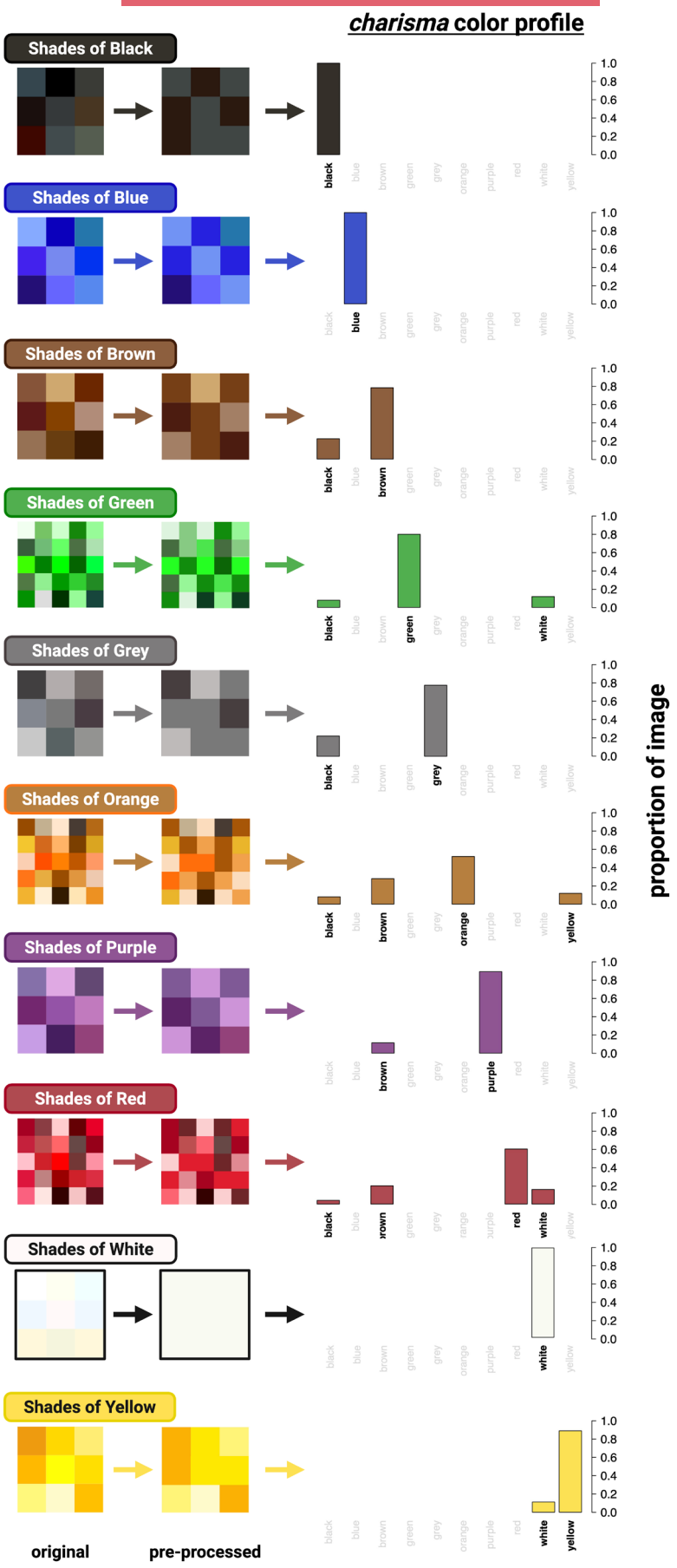

FIGURE 2 Validating `charisma` colour classifications. Swatches, each containing shades of one of ten colour classes, were obtained from Wikipedia (see Table S1) and were then submitted directly to the `charisma` automated workflow. `Charisma` performed well above chance at identifying each of the ten colour classes (see `charisma` colour profile).

TABLE 2 Bird museum specimens.

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

2.5 | Evolutionary analyses

We used our datasets and a previously published tanager phylogeny (Burns et al., 2014) to explore variation in the rates of colour evolution and reconstruct ancestral colour states. Additionally, because the colour-producing mechanisms and structures are well known in birds (Hill, 2006; Mason & Bowie, 2020; Porzio & Mota, 2025; Stoddard & Osorio, 2019; Stoddard & Prum, 2011), we also estimated rates and reconstructed ancestral states for colour-producing mechanisms. We transformed our colour data by grouping discrete colours by avian colour-producing mechanism. We designated melanin-based colours as black, brown and grey, carotenoid-based colours as red, orange and yellow and structural colours (produced by feather nanostructure, including those with pigment overlays) as green, blue and purple (Hill, 2006). For purposes of this study, we have oversimplified the colour-producing mechanisms for green (produced by carotenoids + nanostructure) and blue and purple (both produced by melanin + nanostructure). We removed white from the colour mechanism analysis because it has two potential mechanisms: a lack of pigment or feather nanostructure. For each colour 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 colour are equal, and the all rates different (ARD) model, which assumes gains and losses of a colour occur at different rates. We compared transition rates for every colour in our datasets to determine the potential effects of automated, expert and semi-automated classifications on evolutionary analyses. We used the sample size-corrected Akaike information criterion (AICc) to select the best model for each colour and mechanism for our semi-automated dataset. We then leveraged the best-fitting model to reconstruct ancestral states for our semi-automated dataset. Using the R package `phytools`, we estimated

maximum-likelihood probabilities for each colour 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 colour classification performance

Figure 2 demonstrates `charisma`'s ability to characterize the colours in an image. Here, we find that the majority of *charisma* colour classifications identify the pre-assigned colour class of the input grid. The highest variation in colour classifications was present in the red and orange grids, with four different colours 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 colours in HSV colour space. We demonstrate that diverse sets of organisms can be run using `charisma`, but we warn users that image lighting and quality can change the designation of the colour class depending on the assigned HSV coordinates (see Figure S3 for examples of beetles and fish images). We stress that users should refer to Box 1 for our recommended workflow.

TABLE 3 Performance comparison between the semi-automated **`charisma`** classification workflow (with manual adjustments) and expert classification across colour classes. 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 colour 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.

Colour	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

3.2 | Comparison of colour datasets

We found that the semi-automated colour dataset outperformed automated colour datasets when compared to the expert colour dataset (Table 3; Tables S2–S4). Black was well-classified in all datasets and brown and grey had high false alarm rates in automated datasets (Table 3; Tables S2–S4). The bases and tips of bird feathers, especially contour feathers, are often different colours, 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 & 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 colour artefacts may contribute to the overrepresentation of brown and grey in the automated colour datasets. For our analyses, these patches represented unwanted image artefacts; 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 automated datasets (Table 3; Tables S2–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 automated thresholding procedure removed colours with a small proportion of pixels, making underrepresented colour categories difficult to classify using the automated workflow. Together, the small sample size and small patch size of red and orange in these birds contribute to the high miss rates in the automated datasets. The classifications derived from the semi-automated `charisma` workflow show an almost identical colour profile to the expert colour dataset as evidenced by the near-100% hit and correct rejection rate (Table 3). As such, signal-to-noise ratio (SNR) can be further improved by running the semi-automated mode without thresholding in `charisma` to ensure underrepresented colour patches are called correctly (to reduce misses) and manually correct colour patches sitting at the boundary of the CLUT colour class boundaries (to reduce false alarm identifications).

3.3 | Evolutionary analyses

To understand how classifying colour in different ways (i.e. automated vs. semi-automated datasets) impacts evolutionary rates analyses, we compared evolutionary rates estimated for the expert, semi-automated and automated datasets (Figure 3; Table 4). We found that the automated datasets elevated the rate of evolution of grey across all models providing evidence that uncorrected colour artefacts in bird museum specimens, likely due to misaligned feathers, impact downstream evolutionary analyses. We also found slower evolutionary rates for blue and green in the automated datasets, which may be due to the higher miss rates of `charisma`-identified blue and green colour classes (Table 3; Tables S2–S4). As predicted, we see high congruence in evolutionary rates between our expert and semi-automated colour datasets.

Given the similarity between our expert and semi-automated colour datasets, we present the results of our evolutionary analyses using the best-performing `charisma` classification (i.e. the semi-automated colour dataset). We note that these results and

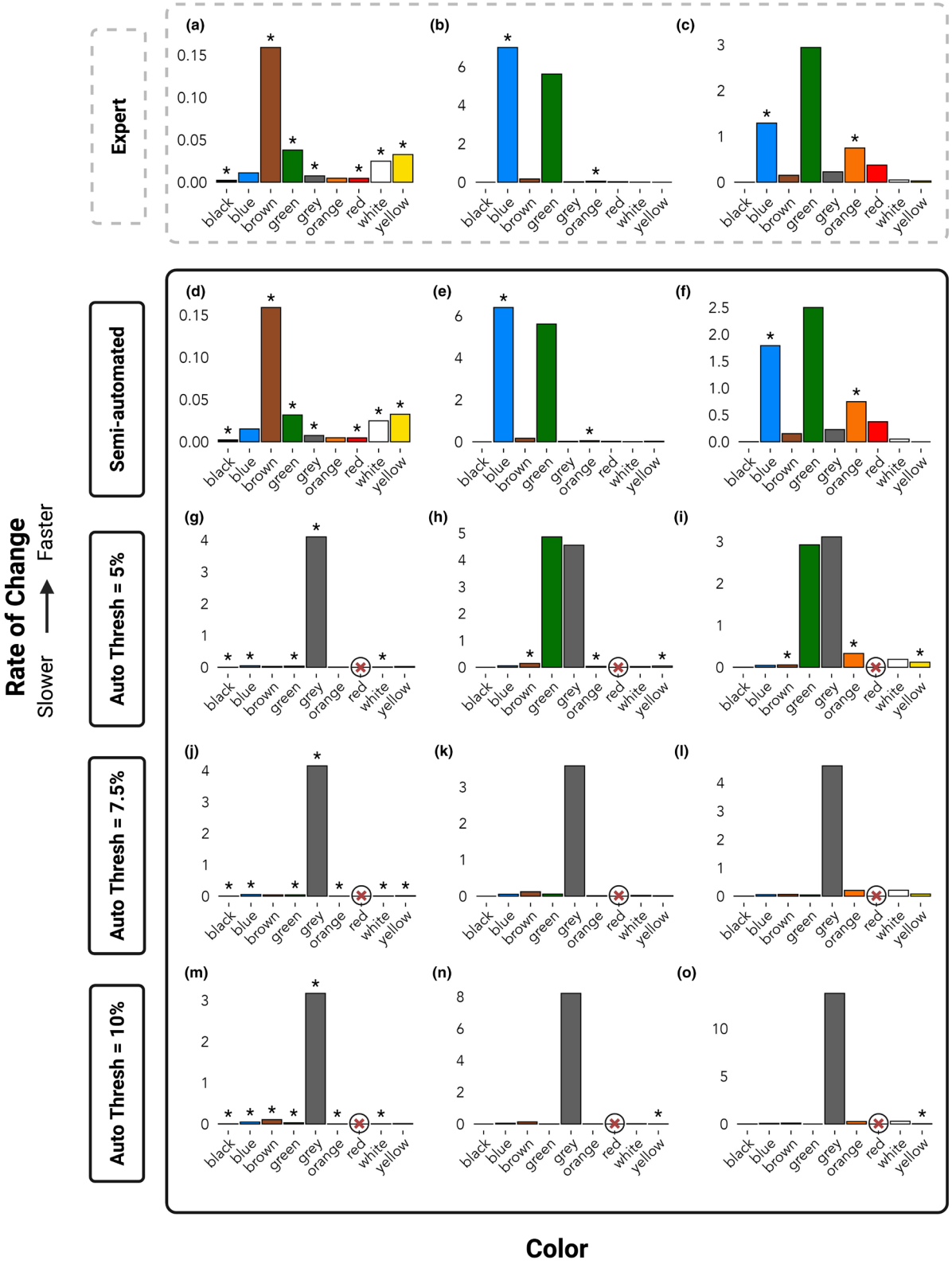

FIGURE 3 Evolutionary rates of change for each colour and evolutionary model as classified in our expert dataset by A.J.S. (a–c) versus `charisma`-based classifications—Semi-automated: Hybrid classification using a threshold = 0% with manual adjustments by W.L.E.T. (d–f) the three remaining classifications were derived from the automated `charisma` workflow with no manual editing, using increasingly conservative colour-proportion thresholds from top to bottom (g–o). Red ⊗ indicates the colour was not called by the `charisma` analysis; asterisks (*) indicate best fit model for each colour class (rates and model selection can be found in Table 4). 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–l), the ER and ARD model selection was inconclusive with an AICc weight of 0.5 for each model.

TABLE 4 Rates of colour evolution and ancestral state reconstruction model selection for all datasets.

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

[a]Best fit model for each colour class.

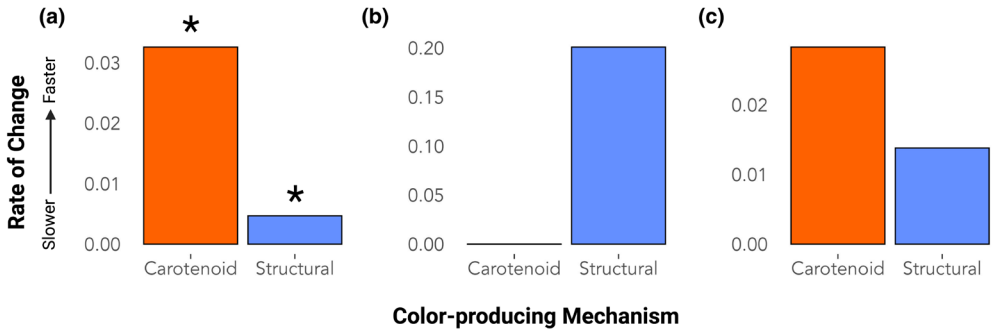

FIGURE 4 Evolutionary rates for each colour-producing mechanism and evolutionary model for the semi-automated dataset: (a) Equal Rates (ER) rate, (b) All Rates Different (ARD) gain rate, (c) ARD loss rate. Asterisks (*) indicate best fit model for each colour-producing mechanism (rates and model selection can be found in Table 5).

TABLE 5 Rates of colour-producing mechanism evolution and ancestral state reconstruction model selection for the semi-automated 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[a]	41.28	0.616	0	0.0283	42.23	0.384
Structural	0.00468[a]	17.35	0.725	0.201	0.0138	19.29	0.275

[a]Best fit model for each colour class.

interpretations are based on incomplete sampling of a subfamily of tanagers and they are presented as an example of the utility of `charisma` for these types of evolutionary analyses. Additionally, we have categorized highly complicated colour-producing mechanisms into two simple mechanisms. Any biological interpretations of the data should be taken with caution. We tested two models of evolution on discrete colour classes and colours grouped by colour-producing mechanism (see Figures 3 and 4; Tables 4 and 5). We excluded purple and the melanin mechanism from our analyses because purple was not present in any species and melanin was ubiquitous across all species. When comparing the evolution of the melanin-based colour classes using the ER model, brown showed an elevated rate compared to all other colour classes, and black showed the slowest rate (Figure 3). Melanin molecules are structurally robust and are often deposited in the primary wing and tail feathers of birds to increase durability and resist abrasion in these high-use feathers (Bonser, 1995). This aligns with our finding that melanin coloration is highly conserved and indicates the need for structural integrity in bird feathers.

For colour-producing mechanisms, carotenoid coloration showed a higher rate of evolution than structural coloration (Figure 4). Pigment-based colour is more widespread across the avian tree of life than structural colour (Hill & McGraw, 2006). However, across our tanager dataset, structural colours (green and blue) are present in more species than carotenoid colours (red, orange and yellow), which may be contributing to the faster rate of evolution for carotenoid coloration. The structural colour, blue, best fit the ARD model of evolution and showed elevated gain and loss rates compared to all other colours (Figure 3; Table 5). Where structural colour is present, there is evidence that colour diversity accumulates faster than pigment-based colour because of the modularity of layering pigments and structure (Eliason et al., 2015; Maia et al., 2016). These subtle structural changes at the nanometre level may result in the production of significantly different colours allowing for rapid evolution and accumulation of colour diversity.

Lastly, we found the ER model of evolution to be the best-fitting model for every colour and colour-producing mechanism except for blue and orange (Table 4). Using the best-fitting model for each colour and mechanism, we reconstructed ancestral states and mapped them across the tanager phylogeny (Figure 5). We found that most ancestral nodes had black, brown, yellow, green and blue colour states, which were also the most common colours in our dataset. Black, white, brown and grey have been shown to be the most common colours across the bird tree of life (Delhey et al., 2023) and differences in our findings are likely due to the uniquely colourful nature of birds in the subfamily Thraupinae. The ancestral state reconstructions demonstrate that rates of evolution are largely driven by losses of colours across internal branches of the tree. Orange fit the ARD model best with a slightly elevated rate of gain than loss (Figure 3). However, orange is only present in one species in our dataset and this result is driven mainly by the uncertainty of the presence of orange at the root of the phylogeny (see Figures 3 and 5).

4 | DISCUSSION

Here, we present `charisma`, an R package we developed to provide a standardized, reproducible and flexible framework for characterizing colour classes from digital images. To demonstrate its utility, we used standardized images of bird museum specimens

BRITISH ECOLOGICAL SOCIETY Methods in Ecology and Evolution

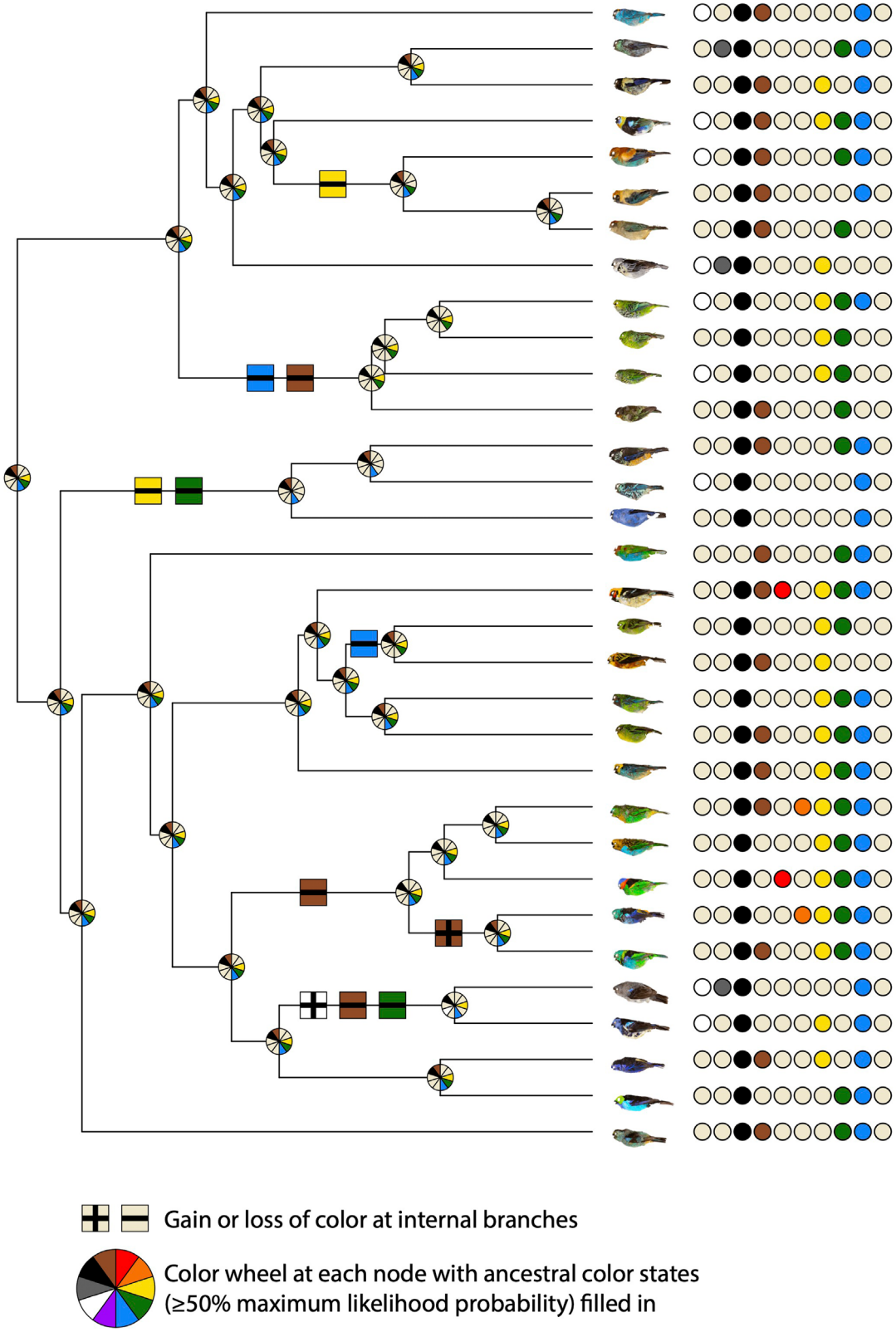

FIGURE 5 `charisma` colour classifications mapped on the tanager phylogeny. For each species, coloured dots represent the colours present in the image. Colour wheels at each node indicate estimated colours with greater than or equal to 50% maximum-likelihood probability. Internal branches are marked with gains or losses of colours. Species order (top to bottom): *Chalcothraupis ruficervix*, *Stilpnia heinei*, *S. cyanoptera*, *S. larvata*, *S. preciosa*, *S. cayana*, *S. vitriolina*, *Poecilostreptus 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. inornata*, *T. mexicana*, *T. velia*, *T. chilensis* and *T. labradorides*.

and found that a hybrid approach using both the automated and semi-automated workflows with the CLUT editor provided the most accurate colour profiling. Artefacts from misaligned feathers in our bird museum specimen image dataset led to overrepresentation of brown and grey in the automated `charisma`-derived colour classification. While these artefacts might not generalize to all image datasets, we recommend using `charisma` in two stages: first, use the automated mode to check and optimize the CLUT classification performance; then, use the semi-automated mode to adjust the outputs using `charisma`'s built-in utilities for classification clean-up.

Overall, `charisma` is built with a fully open source design philosophy where flexibility and customization are both welcomed and encouraged. One example is the built-in functionality to adjust the default CLUT based on user needs and their specific image dataset using the CLUT editor app. Cases where this might be most useful are when fine-tuning is needed for standardized datasets of other organisms, like J.E. Randall's images of fish (http://pbs.bishopmuseum.org/images/JER/) or for extant datasets comprised of non-standardized images, like those culled from iNaturalist (Matheson, 2014; https://www.inaturalist.org) or eBird (Sullivan et al., 2009). We note that image quality and variation can drastically impact downstream analyses and considering them is critical for accurate analyses of colour evolution. We warn users against blindly using the default CLUT without carefully checking colour classifications and refer users to our use recommendations in Box 1.

Lastly, `charisma` enables seamless transition to downstream evolutionary analyses, which we showcased in a vignette on tanagers in the subfamily Thraupinae. More broadly, we present `charisma` as a solution for high-throughput organismal colour analyses with utility that is not restricted to evolutionary biology applications.

AUTHOR CONTRIBUTIONS

Shawn T. Schwartz: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing—original draft preparation; writing—review and editing. **Whitney L. E. Tsai**: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; supervision; validation; visualization; writing—original draft preparation; writing—review and editing. **Elizabeth A. Karan**: Conceptualization; methodology; writing—review and editing. **Mark S. Juhn**: Conceptualization; methodology; writing—review and editing. **Allison J. Shultz**: Methodology; resources; writing—review and editing. **John E. McCormack**: Resources; writing—review and editing. **Thomas B. Smith**: Methodology; writing—review and editing. **Michael E. Alfaro**: Conceptualization; methodology, writing—review and editing.

ACKNOWLEDGEMENTS

We extend our gratitude to current and former members of the Alfaro Lab, Center for Tropical Research and Moore Laboratory of Zoology whose feedback greatly influenced the development of `charisma`. Specifically, we thank Mackenzie Perillo and Kevin Wang for their help in preparing the testing and validating datasets for `charisma`, Gregory Grether for helpful comments on earlier versions of this manuscript and Hannah Weller for insightful and inspiring conversations about colour pattern evolution, computational methods for classifying colour and the development of both the `colordistance` and `recolorize` R packages, of which `charisma` was inspired by and builds upon. Specimens were provided by the Natural History Museum of Los Angeles County.

FUNDING INFORMATION

S.T.S. and W.L.E.T. were supported by the University of California, Los Angeles (UCLA), Department of Ecology and Evolutionary Biology (EEB). W.L.E.T. was additionally supported by the UCLA, Dissertation Year Fellowship; UCLA, EEB, Lida Scott Brown Fellowship; the National Science Foundation, Graduate Research Fellowship (DGE-2034835); and the National Science Foundation, Postdoctoral Research Fellowship in Biology (DBI-2410072). The funders had no role in this project's development, design and delivery.

CONFLICT OF INTEREST STATEMENT

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

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210x.70310.

DATA AVAILABILITY STATEMENT

Data and code available via: Open Science Framework https://osf.io/cqg59 (Schwartz & Tsai, 2025). `charisma` package information available via: R package on CRAN https://cran.r-project.org/package=charisma, GitHub repository https://github.com/shawntz/charisma (Schwartz et al., 2026) and package documents and interactive tutorials https://charisma.shawnschwartz.com.

ORCID

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

REFERENCES

Akkaynak, D., Treibitz, T., Xiao, B., Gürkan, U. A., Allen, J. J., Demirci, U., & Hanlon, R. T. (2014). Use of commercial off-the-shelf digital cameras for scientific data acquisition and scene-specific color calibration. *Journal of the Optical Society of America A*, *31*(2), 312–321.

Alfaro, M. E., Karan, E. A., Schwartz, S. T., & Shultz, A. J. (2019). The evolution of color pattern in butterflyfishes (Chaetodontidae). *Integrative and Comparative Biology*, *59*(3), 604–615.

Andersson, S., & Amundsen, T. (1997). Ultraviolet colour vision and ornamentation in bluethroats. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *264*(1388), 1587–1591.

Barlow, A., Cahill, J. A., Hartmann, S., Theunert, C., Xenikoudakis, G., Fortes, G. G., Paijmans, J. L., Rabeder, G., Frischauf, C., Grandal-d'Anglade, A., García-Vázquez, A., Murtskhvaladze, M., Saarma, U., Anijalg, P., Skrbinšek, T., Bertorelle, G., Gasparian, B., Bar-Oz, G., Pinhasi, R., … Hofreiter, M. (2018). Partial genomic survival of cave bears in living brown bears. *Nature Ecology & Evolution*, *2*(10), 1563–1570.

Bonser, R. H. (1995). Melanin and the abrasion resistance of feathers. *Condor*, *97*(2), 590–591.

Burnham, K. P., Anderson, D. R., Burnham, K. P., & Anderson, D. R. (1998). *Practical use of the information-theoretic approach*. Springer.

Burns, K. J., Shultz, A. J., Title, P. O., Mason, N. A., Barker, F. K., Klicka, J., Lanyon, S. M., & Lovette, I. J. (2014). Phylogenetics and diversification of tanagers (Passeriformes: Thraupidae), the largest radiation of neotropical songbirds. *Molecular Phylogenetics and Evolution*, *75*, 41–77.

Burns, K. J., Unitt, P., & Mason, N. A. (2016). A genus-level classification of the family Thraupidae (Class Aves: Order Passeriformes). *Zootaxa*, *4088*(3), 329–354.

Chan, I. Z., Stevens, M., & Todd, P. A. (2019). Pat-geom: A software package for the analysis of animal patterns. *Methods in Ecology and Evolution*, *10*(4), 591–600.

Commission Internationale de l'Éclairage (CIE). (2004). *Colorimetry. CIE Publication 15:2004*. CIE Central Bureau.

Cooney, C. R., He, Y., Varley, Z. K., Nouri, L. O., Moody, C. J., Jardine, M. D., Liker, A., Székely, T., & Thomas, G. H. (2022). Latitudinal gradients in avian colourfulness. *Nature Ecology & Evolution*, *6*(5), 622–629.

Cooney, C. R., Varley, Z. K., Nouri, L. O., Moody, C. J., Jardine, M. D., & Thomas, G. H. (2019). Sexual selection predicts the rate and direction of colour divergence in a large avian radiation. *Nature Communications*, *10*(1), 1–9.

Dale, J., Dey, C. J., Delhey, K., Kempenaers, B., & Valcu, M. (2015). The effects of life history and sexual selection on male and female plumage colouration. *Nature*, *527*(7578), 367–370.

Darwin, C. (1981). *The descent of man, and selections in relation to sex*. Princeton, NJ: Princeton University Press. (Original work published in 1871).

Delhey, K., Valcu, M., Muck, C., Dale, J., & Kempenaers, B. (2023). Evolutionary predictors of the specific colors of birds. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(34), e2217692120.

Ehrlich, P., Talbot, F., Russell, B., & Anderson, G. (1977). The behaviour of chaetodontid fishes with special reference to lorenz "poster colouration" hypothesis. *Journal of Zoology*, *183*(2), 213–228.

Eliason, C. M., Maia, R., & Shawkey, M. D. (2015). Modular color evolution facilitated by a complex nanostructure in birds. *Evolution*, *69*(2), 357–367.

Endler, J. A. (2012). A framework for analysing colour pattern geometry: Adjacent colours. *Biological Journal of the Linnean Society*, *107*(2), 233–253.

Endler, J. A., Cole, G. L., & Kranz, A. M. (2018). Boundary strength analysis: Combining colour pattern geometry and coloured patch visual properties for use in predicting behaviour and fitness. *Methods in Ecology and Evolution*, *9*(12), 2334–2348.

Feller, K. D., Jordan, T. M., Wilby, D., & Roberts, N. W. (2017). Selection of the intrinsic polarization properties of animal optical materials creates enhanced structural reflectivity and camouflage. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *372*(1724), 20160336.

Garcia, J. E., Girard, M. B., Kasumovic, M., Petersen, P., Wilksch, P. A., & Dyer, A. G. (2015). Differentiating biological colours with few and many sensors: Spectral reconstruction with RGB and hyperspectral cameras. *PLoS One*, *10*(5), e0125817.

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., & Challenger, W. (2008). Geiger: Investigating evolutionary radiations. *Bioinformatics*, *24*(1), 129–131.

Håstad, O., Victorsson, J., & Ödeen, A. (2005). Differences in color vision make passerines less conspicuous in the eyes of their predators. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(18), 6391–6394.

Hemingson, C. R., Cowman, P. F., & Bellwood, D. R. (2024). Analysing biological colour patterns from digital images: An introduction to the current toolbox. *Ecology and Evolution*, *14*(3), e11045.

Hemingson, C. R., Cowman, P. F., Hodge, J. R., & Bellwood, D. R. (2019). Colour pattern divergence in reef fish species is rapid and driven by both range overlap and symmetry. *Ecology Letters*, *22*(1), 190–199.

Hill, G. E. (2006). *Bird coloration, volume 2: Function and evolution*. Harvard University Press.

Hill, G. E., & McGraw, K. J. (2006). *Bird coloration, volume 1: Mechanisms and measurements, volume 1*. Harvard University Press.

Hogan, B. G., & Stoddard, M. C. (2024). Hyperspectral imaging in animal coloration research: A user-friendly pipeline for image generation, analysis, and integration with 3d modeling. *PLoS Biology*, *22*(12), e3002867.

Huelsenbeck, J. P., Nielsen, R., & Bollback, J. P. (2003). Stochastic mapping of morphological characters. *Systematic Biology*, *52*(2), 131–158.

Ibáñez-Álamo, J. D., Delhey, K., Izquierdo, L., Valcu, M., & Kempenaers, B. (2025). Colourful urban birds: Bird species successful in urban environments have more elaborate colours and less brown. *Ecology Letters*, *28*(4), e70106.

Irwin, R. E. (1994). The evolution of plumage dichromatism in the new world blackbirds: Social selection on female brightness. *The American Naturalist*, *144*(6), 890–907.

Karan, E., Schwartz, S., Perillo, M., & Alfaro, M. (2021). It's not just a phase: Evolutionary and functional consequences of sexually dimorphic color pattern diversity in labrid fishes. In *Integrative and Comparative Biology* (Vol. *61*, pp. E442–E443). Oxford University Press.

Karan, E. A., Hodge, J. R., & Alfaro, M. E. (2025). Eyespot function in butterflyfishes (Chaetodontidae): Comparing evidence for aposematism, automimicry, and predator mimicry. *Biological Journal of the Linnean Society*, *144*(2), blae122.

Maia, R., Eliason, C. M., Bitton, P.-P., Doucet, S. M., & Shawkey, M. D. (2013). pavo: An r package for the analysis, visualization and organization of spectral data. *Methods in Ecology and Evolution*, *4*(10), 906–913.

Maia, R., Gruson, H., Endler, J. A., & White, T. E. (2019). pavo 2: New tools for the spectral and spatial analysis of colour in r. *Methods in Ecology and Evolution*, *10*(7), 1097–1107.

Maia, R., Rubenstein, D. R., & Shawkey, M. D. (2016). Selection, constraint, and the evolution of coloration in African starlings. *Evolution*, *70*(5), 1064–1079.

Mason, N. A., & Bowie, R. C. (2020). Plumage patterns: Ecological functions, evolutionary origins, and advances in quantification. *The Auk*, *137*(4), ukaa060.

Matheson, C. A. (2014). Inaturalist. *Reference Reviews*, *28*(8), 36–38.

Mayr, E. (1963). *Animal species and evolution*. Harvard University Press.

McMillan, W. O., Weigt, L. A., & Palumbi, S. R. (1999). Color pattern evolution, assortative mating, and genetic differentiation in brightly colored butterflyfishes (Chaetodontidae). *Evolution*, *53*(1), 247–260.

Neudecker, S. (1989). Eye camouflage and false eyespots: Chaetodontid responses to predators. In M. Sano (Ed.), *The butterflyfishes: Success on the coral reef* (pp. 143–158). Springer.

Nicholson, K. E., Harmon, L. J., & Losos, J. B. (2007). Evolution of *Anolis* lizard dewlap diversity. *PLoS One*, *2*(3), e274.

Paradis, E., Claude, J., & Strimmer, K. (2004). Ape: Analyses of phylogenetics and evolution in r language. *Bioinformatics*, *20*(2), 289–290.

Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., Alfaro, M. E., & Harmon, L. J. (2014). geiger v2. 0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, *30*(15), 2216–2218.

Porzio, N. S., & Mota, P. G. (2025). Evidence for the evolution of male and female cardinals' plumage colouration being affected by both natural and sexual selection. *BMC Ecology and Evolution*, *25*(1), 1–11.

Powers, D. M. (2020). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*. https://doi.org/10.48550/arXiv.2010.16061

Price-Waldman, R. M., Ali, J. R., Shultz, A. J., Hogan, B. G., & Stoddard, M. C. (2025). Hidden white and black feather layers enhance plumage coloration in tanagers and other songbirds. *Science Advances*, *11*(30), eadw5857.

Price-Waldman, R. M., Shultz, A. J., & Burns, K. J. (2020). Speciation rates are correlated with changes in plumage color complexity in the largest family of songbirds. *Evolution*, *74*(6), 1155–1169.

Rabosky, A. R. D., Cox, C. L., Rabosky, D. L., Title, P. O., Holmes, I. A., Feldman, A., & McGuire, J. A. (2016). Coral snakes predict the evolution of mimicry across new world snakes. *Nature Communications*, *7*(1), 1–9.

Revell, L. J. (2012). phytools: An r package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *2*, 217–223.

Revell, L. J. (2024). phytools 2.0: An updated r ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, *12*, e16505.

Santangeli, A., Haukka, A., Morris, W., Arkkila, S., Delhey, K., Kempenaers, B., Valcu, M., Dale, J., Lehikoinen, A., & Mammola, S. (2023). What drives our aesthetic attraction to birds? *npj Biodiversity*, *2*(1), 20.

Schwartz, S., Tsai, W., Karan, E., Juhn, M., Shultz, A., McCormack, J., Smith, T., & Alfaro, M. (2026). shawntz/charisma: Charisma v1.1.0. *Zenodo*. https://doi.org/10.5281/zenodo.19491931

Schwartz, S. T., & Alfaro, M. E. (2021). Sashimi: A toolkit for facilitating high-throughput organismal image segmentation using deep learning. *Methods in Ecology and Evolution*, *12*(12), 2341–2354.

Schwartz, S. T., & Tsai, W. L. (2025). *charisma: An r package to perform reproducible color characterization of digital images for biological studies*. OSF. https://osf.io/cqg59

Senior, R. A., Oliveira, B. F., Dale, J., & Scheffers, B. R. (2022). Wildlife trade targets colorful birds and threatens the aesthetic value of nature. *Current Biology*, *32*(19), 4299–4305.

Shultz, A. J., & Burns, K. J. (2013). Plumage evolution in relation to light environment in a novel clade of neotropical tanagers. *Molecular Phylogenetics and Evolution*, *66*(1), 112–125.

Shultz, A. J., & Burns, K. J. (2017). The role of sexual and natural selection in shaping patterns of sexual dichromatism in the largest family of songbirds (Aves: Thraupidae). *Evolution*, *71*(4), 1061–1074.

Stevens, M., Párraga, C. A., Cuthill, I. C., Partridge, J. C., & Troscianko, T. S. (2007). Using digital photography to study animal coloration. *Biological Journal of the Linnean Society*, *90*(2), 211–237.

Stoddard, M. C., Eyster, H. N., Hogan, B. G., Morris, D. H., Soucy, E. R., & Inouye, D. W. (2020). Wild hummingbirds discriminate nonspectral colors. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(26), 15112–15122.

Stoddard, M. C., & Osorio, D. (2019). Animal coloration patterns: Linking spatial vision to quantitative analysis. *The American Naturalist*, *193*(2), 164–186.

Stoddard, M. C., & Prum, R. O. (2011). How colorful are birds? Evolution of the avian plumage color gamut. *Behavioral Ecology*, *22*(5), 1042–1052.

Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). ebird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, *142*(10), 2282–2292.

Terrill, R. S., & Shultz, A. J. (2023). Feather function and the evolution of birds. *Biological Reviews*, *98*(2), 540–566.

Valvo, J. J., Rodd, F. H., Aponte, J. D., Daniel, M., Dwinell, K., Houle, D., & Hughes, K. A. (2020). Colormesh: A novel method for quantifying variation in complex color patterns. *bioRxiv*. https://doi.org/10.1101/2020.07.17.205369

Van Belleghem, S. M., Papa, R., Ortiz-Zuazaga, H., Hendrickx, F., Jiggins, C. D., Owen McMillan, W., & Counterman, B. A. (2018). patternize: An r package for quantifying colour pattern variation. *Methods in Ecology and Evolution*, *9*(2), 390–398.

van den Berg, C. P., Troscianko, J., Endler, J. A., Marshall, N. J., & Cheney, K. L. (2020). Quantitative colour pattern analysis (QCPA): A comprehensive framework for the analysis of colour patterns in nature. *Methods in Ecology and Evolution*, *11*(2), 316–332.

Weller, H. I., Hiller, A. E., Lord, N. P., & Van Belleghem, S. M. (2024). recolorize: An r package for flexible colour segmentation of biological images. *Ecology Letters*, *27*(2), e14378.

Weller, H. I., & Westneat, M. W. (2019). Quantitative color profiling of digital images with earth mover's distance using the r package colordistance. *PeerJ*, *7*, e6398.

Yu, J., Duan, H., Zhang, B., Zhang, L., & He, J. (2024). Urbanization alters the geographic patterns of passerine plumage color in China. *Landscape and Urban Planning*, *248*, 105101.

SUPPORTING INFORMATION

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

Figure S1. Full HSV colour space (a) and colour slices for each colour class (b) as visualized using the CLUT editor.

Figure S2. Comparison of original image (a), white-balanced image (b) and white-balanced and segmented image (c) for *Tangara fastuosa*.

Figure S3. `charisma` colour classifications for fish (Alfaro et al., 2019; Karan et al., 2021, 2025; Schwartz & Alfaro, 2021) and beetles (Weller et al., 2024).

Table S1. Sources images for Wikipedia-based shades of colour swatch validation.

Table S2. Performance comparison between the automated `charisma` classification workflow (with a 5% colour-proportion threshold) and expert classification across colour classes.

Table S3. Performance comparison between the automated `charisma` classification workflow (with a 7.5% colour-proportion threshold) and expert classification across colour classes.

Table S4. Performance comparison between the automated `charisma` classification workflow (with a 10% colour-proportion threshold) and expert classification across colour classes.

How to cite this article: Schwartz, S. T., Tsai, W. L. E., Karan, E. A., Juhn, M. S., Shultz, A. J., McCormack, J. E., Smith, T. B., & Alfaro, M. E. (2026). **`charisma`**: An R package to perform reproducible colour characterization of digital images for biological studies. *Methods in Ecology and Evolution*, *17*, 1703–1718. https://doi.org/10.1111/2041-210x.70310