

Evolution of correlated complexity in the radically different courtship signals of birds-of-paradise

5 Russell A. Ligon^{1,2*}, Christopher D. Diaz¹, Janelle L. Morano¹, Jolyon Troscianko³, Martin Stevens³,
Annalyse Moskeland^{1□}, Timothy G. Laman⁴, Edwin Scholes III¹

1- Cornell Lab of Ornithology, Ithaca, USA.

10 2- Department of Neurobiology and Behavior, Cornell University, Ithaca, USA.

3- Centre for Ecology and Conservation, College of Life and Environmental Science, University of Exeter, Penryn, UK

4- Museum of Comparative Zoology, Harvard University, Cambridge, USA

15 * russell.ligon@gmail.com

□ Current address: Department of Zoology, University of Oxford, UK

ORCID:

Russell Ligon <https://orcid.org/0000-0002-0195-8275>

20 Janelle Morano <https://orcid.org/0000-0001-5950-3313>

Edwin Scholes <https://orcid.org/0000-0001-7724-3201>

Jolyon Troscianko <https://orcid.org/0000-0001-9071-2594>

25 russell.ligon@gmail.com

cdd67@cornell.edu

janelle.morano@cornell.edu

jt@jolyon.co.uk

Martin.Stevens@exeter.ac.uk

30 annalysemoskeland@gmail.com

edwin.scholes@cornell.edu

keywords: ornament, complexity, behavioral analyses, sensory ecology, phenotypic radiation

35

40

ABSTRACT

Ornaments used in courtship often vary wildly among species, reflecting the evolutionary interplay between mate preference functions and the constraints imposed by natural selection. Consequently, understanding the evolutionary dynamics responsible for ornament diversification has been a longstanding challenge in evolutionary biology. However, comparing radically different ornaments across species, as well as different classes of ornaments within species, is a profound challenge to understanding diversification of sexual signals. Using novel methods and a unique natural history dataset, we explore evolutionary patterns of ornament evolution in a group – the birds-of-paradise – exhibiting dramatic phenotypic diversification widely assumed to be driven by sexual selection. Rather than the trade-off between ornament types originally envisioned by Darwin and Wallace, we found positive correlations among cross-modal (visual/acoustic) signals indicating functional integration of ornamental traits into a composite unit – the *courtship phenotype*. Furthermore, given the broad theoretical and empirical support for the idea that systemic robustness – functional overlap and interdependency – promotes evolutionary innovation, we posit that birds-of-paradise have radiated extensively through ornamental phenotype space as a consequence of the robustness in the courtship phenotype that we document at a phylogenetic scale. We suggest that the degree of robustness in courtship phenotypes among taxa can provide new insights into the relative influence of sexual and natural selection on phenotypic radiations.

60

65

70

75 **Author Summary**

Animals frequently vary widely in ornamentation, even among closely related species. Understanding the patterns that underlie this variation is a significant challenge, requiring comparisons among drastically different traits – like comparing apples to oranges. Here, we use novel analytical approaches to quantify variation in ornamental diversity and richness across the wildly divergent birds-of-paradise, a textbook example of how sexual selection can profoundly shape organismal phenotypes. We find that color and acoustic complexity, along with behavior and acoustic complexity, are positively correlated across evolutionary time-scales. Positive links among ornament classes suggests that selection is acting on correlated suites of traits – a composite courtship phenotype – and this integration may be partially responsible for the extreme variation in signal form that we see in birds-of-paradise.

85

Adaptive radiations are driven by ecological differences that promote processes of diversification and speciation [1]. In contrast, phenotypic radiations which occur in the absence of clear ecological differentiation are less well-understood. One commonly investigated mechanism for phenotypic diversification among ecologically similar taxa is variation in social and sexual selection pressures promoting signal or ornament diversification. Ornamental radiations may come about as a consequence of variation in signaling environment [2,3], sensory capabilities [4,5], or pseudo-randomly via mutation-order selection [6,7] or Fisher-Lande-Kirkpatrick processes [8–11]. Most studies investigating patterns of ornamental diversification have focused on individual trait classes and simplified axes of variation, however, sexual selection does not act on single traits in isolation. A more complete understanding of the processes driving ornamental diversification is possible only by investigating evolutionary relationships between the full suites of ornamental traits under selection.

Many animals rely on multiple ornamental traits to attract mates. Advantages of multiple ornaments may include increased information transfer (multiple messages), increased reliability (redundancy), increased flexibility (ensuring information transfer across contexts and environments), and increased memorability/discriminability [12–16]. Multiple ornaments may be more common when costs associated with the display or evaluation of those ornaments is low [17], as is likely the case in lekking species [18]. Though we now have broad empirical support for many of the proposed adaptive benefits of multiple signals at the level of individual species, how these specific hypotheses map onto our understanding of phylogenetic patterns of ornament evolution is less clear. Insights into the macroevolutionary patterns of multiple ornament evolution are challenging, in part, owing to the difficulties of comparing highly divergent phenotypic traits across species. For example, even focusing on evolutionary patterns of a single trait (e.g. plumage color in birds) across species can be difficult when traits possess different axes of variation (e.g. red vs. blue). Though ingenious new methods have been devised to compare highly divergent ornaments of a single signal type (e.g. plumage color [19], electrical signals [20], or song [21]), comparing ornamental complexity across signal types presents yet an additional layer of complication. However, understanding the interrelationships of different classes of ornaments across phylogenetic scales can potentially provide valuable information about the evolutionary processes of communication, phenotypic radiation, and speciation that cannot be gathered from single trait or single species studies.

Following the evolution of multiple ornaments, selective pressures may favor different interrelationships among signal types. If ornamental investment is governed by evolutionary trade-offs, investment in one class of ornaments will come only at the expense of investment in another. Evidence suggests that signal trade-offs manifests as a negative correlation among ornament types across evolutionary time [22–28], reflecting strong, consistent constraints imposed by ecology, physiology, and natural selection [29,30]. Alternatively, instances where ornamental traits show no evolutionary relationships [31–36] suggest long-term patterns of independent evolutionary trajectories. In such cases, signals are functionally independent and may even have evolved for use in different contexts (e.g. territorial defense vs. mate attraction). When might we expect positive correlations among ornament classes across species? Theoretical [18,37] and empirical [38,39] work suggest that positive correlations among signals across species may reflect consistent selection acting similarly on separate axes of ornamental evolution. Strong, consistent intersexual selection could generate these positive correlations (*sensu* [38]), especially if the signals convey separate information [37], resulting in functional integration among ornament elements [40,41]. In such cases, positive correlations among signals across species would arise when selection favors an ‘integrated whole’ of ornamental traits [42,43], which we call the *courtship phenotype*. The courtship phenotype is the composite expression of all ornamental classes evaluated during courtship and may represent the composite target of selection. Evolution may favor integrated, holistic mate evaluation strategies because of advantages that sensory overlap and redundancy offer (e.g. increased accuracy) [12–16].

Here, we examine broad evolutionary patterns of ornamental signal investment and complexity across the wildly diverse [44,45], monophyletic [46] birds-of-paradise (Paradisaeidae) “...in which the process of sexual selection has gone to fantastic extremes” [47] (Fig 1A). We focus on the birds-of-paradise because this family exhibits extreme variation across species in multiple ornamental axes [44] (e.g. color [48–50], behavior [51,52]), while possessing broadly similar life-histories and mating systems [44,53]. Consequently, insights about the strength, direction, and diversification of ornamental phenotypes in this group may shed light on key processes of sexual selection and its power to generate phenotypic radiation when natural selection-imposed constraints are minimized. In this study, we use a unique natural history dataset to quantitatively evaluate behavioral, acoustic, and colorimetric ornamentation across 40 species of birds-of-paradise, as well as relationships between signals and display environment.

Fig 1. Birds-of-paradise exhibit extreme diversity in colors, sounds, and behaviors used during courtship displays, necessitating novel methods to quantitatively evaluate the evolution of their complexity. (A) Sixteen exemplar species (purple tips) are shown with their phylogenetic relationships to highlight variation in plumage color, acoustic signals, and courtship display behavior. (B) Behavioral sub-units were scored from field-captured videos of displaying males (S3, S4 Table). Behavioral sub-units were combined to create composite behaviors describing any behavior, across species, and facilitating sliding-window analysis of behaviors and behavioral sequences. (C) Ultraviolet and visual spectrum images were taken of museum specimens (S7 Table) and used to generate avian visual model-informed image stacks. Color values were clustered with respect to modeled avian discriminability, enabling whole-specimen quantification of color richness and diversity. (D) All bird-of-paradise sounds were placed into a multidimensional acoustic space defined by principal components analysis. Sounds were then given identities based on locations within acoustic-space, facilitating a sliding-window analysis of sounds and acoustic sequences (S8, S9 Table).

RESULTS

An approach to quantify courtship complexity among divergent ornaments

Comparisons across signal types are inherently challenging for evolutionary biologists given that such signals are necessarily measured in different ways. Additionally, comparisons within color, acoustic, and behavioral repertoires across taxa that vary widely (e.g. the birds-of-paradise) present an additional methodological challenge: how does one compare phenotypes that may share no obvious overlapping characters? We addressed this obstacle with a two-pronged approach to quantify ornamental complexity for behavior, color, and sounds in the birds-of-paradise. First, we broke down each ornament into a taxonomically-unbounded character space that allowed classification of subunits across all species. Second, we used the specific attributes of a given ornament for each individual, for each species, to categorize the ornament components before quantifying two conceptually-aligned measures of complexity for each signal type. Specifically, we evaluated richness (the number of unique elements) and diversity (using an index dependent on the number and relative contribution of each element type) using phylogenetic comparative approaches (see *Methods* for additional details).

For behavioral analyses, we first broke down the courtship behaviors of all species into distinct sub-units, shared across species (e.g. [S1 Video 1](#)). We then analyzed composite behavioral sequences across time using sliding-window analyses to compare maximally diverse behavioral repertoires for a set duration across species (Fig 1B). For colorimetric analyses, we relied on visual modeling of multispectral images to quantify the number and relative abundances of perceptually-distinct color types across individuals and species. Though different colors may have different underlying production mechanisms, our analyses simply focused on the number and distribution of distinguishable colors (Fig 1C). Similar to our behavioral analysis pipeline, we used acoustic properties and agglomerative clustering to classify distinct sound-types used by bird-of-paradise in courtship contexts before employing a similar sliding-window analysis to identify maximally diverse acoustic sequences, facilitating comparisons across species (Fig 1D).

195 In total, we analyzed 961 video clips, 176 audio clips, and 393 museum specimens. From these analyses, we obtained quantitative diversity and richness metrics of ornamental complexity across the birds-of-paradise (Fig 2), which allowed us to rigorously evaluate patterns of correlated character evolution, as well as facilitating our investigation of the influence of breeding system and display environment on ornamental complexity.

200 **Fig 2. Signal diversity indices vary widely across birds-of-paradise.** To facilitate interpretation of variation in signal diversity across the family (Paradisaeidae), we generated scaled diversity scores for each signal type (color, behavior, sound) to create (left) a composite metric of overall signal diversity (heat-mapped to the phylogeny) and (right) stacked bar plots illustrating variation in the relative diversity for each signal type for each species.

205 **Integrative evolution of courtship complexity across modalities**

205 Using multiple phylogenetic generalized least squares analyses, which allowed us to control for the non-independence of species due to their shared evolutionary history, as well as the potentially confounding influences of display environment (both display height and proximity to courting conspecific males), we uncovered positive correlations between color and acoustic diversity (Fig 3A), as well as between behavioral and acoustic diversity (Fig 3B), consistent with the hypothesis that selection has acted similarly on these axes of ornamental complexity. Interestingly, however, there was no significant relationship between color and behavioral diversity, indicating independent evolutionary trajectories for these visually-encoded aspects of courtship ornamentation (S1 Table, S1 Fig). Analyses of ornamental richness revealed the same pattern to those uncovered for ornamental diversity (S2 Table). Specifically, behavioral and acoustic richness were correlated, as were color and acoustic richness (as was the case for both relationships involving ornamental diversity).

220 **Fig 3. Positive phylogenetic correlations exist among several ornamental diversity indices at an evolutionary scale.** (A) Color and acoustic diversity are positively correlated in the birds-of-paradise, where species with greater color diversity exhibit increased acoustic diversity when controlling for behavioral diversity, display height, and display proximity in a multiple phylogenetic least squares regression (mPGLS; summary statistics in S1 Table). (B) Behavior and acoustic diversity are positively correlated in the birds-of-paradise, where species with greater behavioral diversity exhibit increased acoustic diversity when controlling for color diversity, display height, and display proximity in a multiple phylogenetic least squares regression (mPGLS; summary statistics in S1 Table). Species' points represent tip values for log transformed behavioral and color diversity. Underlying data for Fig 3 can be found in S1 Data.

230 **Courtship complexity related to display height**

230 Behavior and acoustic, but not color, richness were influenced by stratum of the forest in which species display (Fig 4A-C). Specifically, we found that behavioral richness exhibited a negative relationship with display height among birds-of-paradise, such that species that display on the forest floor had the largest behavioral repertoires (S2 Table, Fig 4B). Species that display on the forest floor are typically operating with lower-light environments, and consequently these species appear to rely more heavily on complex dance sequences to attract mates. Additionally, birds-of-paradise show increased acoustic (Fig 4C) richness as their display locations increase in height (S2 Table), a result that partially corresponds to the predictions of sensory drive [54,55] whereby the openness of the upper-canopy favors increasingly complex acoustic displays.

240 Similar to the patterns we uncovered for signal richness, we also found that behavioral and acoustic diversity were influenced by display height (S1 Table). Species displaying in the forest understory exhibiting a marginally significant ($p = 0.051$) trend for greater acoustic diversity relative to ground displaying species, and the behavioral diversity for ground-displaying species was higher than for both understory and canopy species (S1 Table). However, color diversity was not significantly influenced by display height.

245 **Courtship complexity related to spatial distribution of displaying males**

Birds-of-paradise that display in classic leks have greater color richness (Fig 4D, S2 Table), corresponding to the increased strength of sexual selection on males to ‘stand out’ visually when being evaluated simultaneously in lekking contexts. However, neither behavioral nor acoustic richness were significantly associated with the spatial distribution of displaying males. Furthermore, none of the diversity metrics (color, behavior, sound) were significantly associated with the breeding system structure (S1 Table).

255 **Fig 4. Social and environmental variation in display microhabitat influences multiple axes of ornamental complexity in birds-of-paradise.** Display height did not influence color richness (A) but did influence behavioral (B) and acoustic (C) richness. Specifically, species that display on the forest floor have larger behavioral repertoires than species that display in the understory and canopy (B). Additionally, species that display in the canopy have larger acoustic repertoires compared to ground-displaying species (C). Social dynamics (D-F) of the display environment, measured as the proximity of other courting males, influenced color, but not behavioral or acoustic, richness in the birds-of-paradise. Species that display in classic leks have greater color richness (D), than species that display in exploded-leks or solitarily. However, neither behavioral (E) or acoustic (F) richness were significantly influenced by breeding/display system. Violin plots illustrate the distribution of log-transformed richness scores for each species. Underlying data for Fig 4 can be found in S1 Data.

260 **DISCUSSION**

Our study provides evidence that selection has favored correlated levels of ornamental diversity across multiple signals among the birds-of-paradise. This pattern of positive correlation among distinct ornament classes across evolutionary time-scales and species suggests strong sexual selection on functionally integrated *courtship phenotypes*. The degree to which phenotypic traits are co-expressed and functionally dependent upon one another can be referred to as functional integration [56] or interdependence [41]. Courtship phenotypes with greater functional integration are, therefore, composed of ornaments that are typically expressed at similar levels and which are mutually interdependent in order to influence mate choice [40,42,43]. Correlations among the signals that comprise the courtship phenotype also suggest a previously undescribed robustness in bird-of-paradise courtship phenotypes that may have played a key role in the extreme ornamental radiation exhibited by this taxon (Fig 1).

275 Evolutionary biologists dating back to Mayr [57] and even Darwin [58] have recognized the potential evolutionary implications of functional redundancy (two or more structures performing the same function). Functional redundancy, including ‘true’ redundancy (i.e. structurally identical components with identical functions) and degeneracy (i.e. structurally distinct components with similar functions) [59], facilitates evolutionary innovation (i.e. increases ‘evolvability’) by increasing robustness. Robust systems are those where the overall structure and interconnectedness of parts provides protection from environmental or mutational instability [60] such that a given function is not lost if a single component fails. Robustness increases evolvability by enabling elements to react to selection independently and diverge while maintaining original functions [59,61]. All redundancy (both ‘true’ redundancy and degeneracy) provides a measure of robustness, but robust systems are not necessarily redundant [62]. Given the broad theoretical [61,63,64] and empirical [65–67] support for the idea that robustness can promote evolvability across a wide array of biological domains, we posit that the correlations among signal types within birds-of-paradise courtship phenotypes are, at least partially, responsible for the dramatic diversification and radiation of courtship signals displayed by birds-of-paradise. If female birds-of-paradise make mate choice decisions based on sensory input from the multiple signals that comprise a composite courtship phenotype, and information from those channels is correlated, then novel mutations changing the structure or form of a given ornament may occur without “necessary” information being lost [68]. Consequently, over evolutionary time, we suggest it is the inherent functional overlap (redundancy/degeneracy) and structural interdependency (robustness) of courtship phenotypes that leads to increased phenotypic diversification (evolvability) in birds-of-paradise.

295 Phenotypic radiations in the absence of clear ecological differentiation may arise stochastically [69,70] and
be heavily influenced by the specific intricacies of female choice [7,10,71,72]. Birds-of-paradise clearly
exhibit some ecological differentiation [44] but, broadly speaking, they tend to be heavily frugivorous and
predominantly polygynous [73]. They do not, however, all display to potential mates in the same contexts
300 or microenvironments. Some species display high in the canopy, some down on the forest floor, and others
in the understory in between. Likewise, some species display in large, cacophonous leks, some in exploded
leks where males can hear but not see one another, and other species display solitarily. Our results suggest
that these differences have shaped the specific courtship and signaling strategies of each species (Fig 4, S1-
2 Table). Birds with richer acoustic repertoires display high in the canopy where there is less environmental
305 interference (e.g. from cluttered branches), increasing the likelihood that females will be able to detect and
discern numerous, elaborate sounds [55]. Likewise, more-behaviorally-complex birds tend to display near
the forest floor where there is less light (and ability to perceive subtle variation in color) but more area
available for a courtship stage or “dance floor”. Birds that display in true leks have more colorful plumage,
perhaps because females need to identify attractive individuals based on relatively unchanging traits
310 allowing them to compare among multiple displaying males simultaneously. Display-site and display-
context thus influence the specific forms of ornamentation possessed by individual species [74], and taking
them into account from an analytical perspective allows us to better understand patterns of signal
coevolution and the potential importance of a functionally integrated courtship phenotype.

315 Signal efficacy and information content can exert strong influence on receiver preferences, and
understanding both elements is integral when examining the evolution of complex, multicomponent
courtship phenotypes [14,72,75,76]. The influence of receiver preference is difficult to overstate,
particularly in birds-of-paradise, where recent work indicates that selection acting on female preferences
controls the rate, extent, and phenotypic space available for ornamental radiations [72]. Importantly,
320 receiver preferences are influenced by the perceptual abilities [77,78] and psychology of signal receivers
[79,80], as well as the environments through which signals are transmitted [54] – all of which can markedly
influence signal efficacy. Additionally, the information content of multiple signals may increase the net
amount of information transferred (e.g. multiple messages [16]) or increase accuracy and reliability if
multiple signals communicate the same message (e.g. redundant signals [12,16]). The perceptual channels
325 by which birds-of-paradise attract mates and those channels that are correlated at a phylogenetic scale
provide tantalizing, though tentative, insights into the processes of efficient information transfer and
receiver stimulation regulating mate choice in this group. Specifically, the fact that significant positive
correlations exist between acoustic and color signals (auditory, visual), and between acoustic and behavioral
signals (auditory, visual), but *not* between color and behavioral signals (visual, visual) aligns with
330 psychometric literature on information and sensory input. When multiple sources of information are
provided, information may be maximized if that information comes from separate channels (e.g. acoustic,
visual) and lost when arriving through a single sensory channel [81] (but see [82]). What exactly this
‘information’ might be in birds-of-paradise (quality [83], attractiveness [71], motivation [84], etc.) is not
clear, but this result provides an interesting starting point for future investigations.

335 Phylogenetic comparative investigations of animal signals hold the potential to answer important questions
about the evolutionary trajectories of communication over time [85,86]. However, the data used to tackle
key questions of signal evolution necessarily place upper and lower bounds on the confidence and
interpretations one can make from such comparative studies. It is our hope that the novel approaches we
340 have developed to quantify color, sound, and behavior will be useful to other researchers interested in
understanding signal variation at different scales. Though our primary aim was to generate methodological
pipelines that facilitated comparisons among the highly divergent birds-of-paradise, the basic framework
we describe here may also be useful for comparisons of more similar taxa – including studies of intraspecific
variation in signaling effort (e.g. through sliding-window analyses focused on bouts of maximal
345 complexity) or investment (e.g. by using receiver visual models to identify the number and perceptual
similarity of color patches across individuals). Consequently, we feel that our approaches complement

recent suggestions for incorporating a systems biology approach to the study of animal communication [59] – wherein more-comprehensive, higher resolution data will only improve the validity and interpretability of analyses incorporating fitness surfaces and communication networks.

350 Evolutionary trade-offs – increases in trait expression linked to reductions in another – are ubiquitous: “If
there were no trade-offs, then selection would drive all traits correlated with fitness to limits imposed by
history and design” [87]. Trade-off thinking can inform our interpretations of both the marked interspecific
variation in overall signal complexity (Fig 2) and the finding that the ornaments of birds-of-paradise are
355 positively correlated at phylogenetic scale (Fig 3). Firstly, interspecific variation in overall signal
complexity suggests tradeoffs between investment in courtship and some other, unmeasured, ecological
variable that differs across species (e.g. microenvironment, resource competition, etc.). Secondly, the
absence of trade-offs among signal types suggests an absence of differential costs on acoustic, behavioral,
and chromatic signals. Further, the correlation among ornamental classes suggests that selection is acting
360 on functionally integrated courtship phenotypes for birds-of-paradise, a finding that suggests female birds-
of-paradise make mate choice decisions incorporating holistic, multicomponent information sets comprised
of the various ornaments possessed by males of their species. Rather than being unique to birds-of-paradise,
however, we suggest that this phenomenon is widespread among animals – though is at varying degrees
constrained, impeded, or obfuscated by conflicting and constraining processes and limitations imposed by
ecology and natural selection. The degree to which selection has facilitated the evolution of integrated,
365 robust courtship phenotypes may in fact serve as a proxy for the overall strength and consistency of female-
driven sexual selection in any taxa, where the integration and correlation among ornaments comprising the
courtship phenotype may shed important light on the history and strength of sexual selection in that
particular group.

370 METHODS

Ethics statement

The study was focused on vertebrates (birds-of-paradise) but used museum specimens (physical and media) so no IACUC protocol was required.

375

Behavioral complexity

We quantified the behavioral complexity of courtship display behaviors for the birds-of-paradise by scoring field-recorded video clips of 32 (80%) paradisaeid species, primarily from the Macaulay Library at the Cornell Lab of Ornithology (macaulaylibrary.org, S3 Table). In total, we watched 961 clips from 122 individuals totaling 47707.2 seconds (≈ 795.12 minutes; mean clip duration = 49.64s). Courtship display behavior is highly variable among BOP species, necessitating broad behavioral categories to facilitate investigations of behavioral evolution. Specifically, one of us (CDD) blindly evaluated video clips of male birds-of-paradise displaying species-typical courtship behaviors [44] using a customized ethogram of behavioral units that enabled us to quantify all state and event behaviors exhibited by all species of Paradisaeidae (S4 Table).

380

385

Data collection

To record courtship display behaviors, we used a customized version of an open-source behavior logging program [88]. Additionally, we created a customized keyboard that allowed us to quickly and accurately record the start/stop times of all duration behaviors, as well as the instances of all event behaviors. The combinations of different behavioral categories throughout each clip allowed us to generate sequence data of distinct behavioral elements.

390

Measures of sexual display behavior complexity

Courtship displays can be broken down into distinct behavioral elements and the transitions between these elements. We investigated the number of unique behavioral elements (*behavioral richness*) in a given time period, as well as the Shannon entropy [89] of these behaviors (*behavioral diversity*). Shannon entropy provides a measure of ‘information’ encoded in the behavioral displays, and we converted Shannon entropy scores to their numbers equivalents [90,91]. Shannon indices were chosen specifically because they are the only measures that “give meaningful results when community weights are unequal” [90]. As previously described [90], the numbers equivalent for Shannon entropy values has the readily interpretable property whereby a value of $2x$ would indicate a behavioral sequence with twice as many equally-well-represented behaviors as a sequence with a value of x . In the context of behavioral displays, birds that use many unique behaviors and spend roughly equal amounts of time performing each display element (increased evenness as a proportion of time) will have higher diversity scores.

395

400

405

Sliding window analysis

The number of courtship recordings available was highly variable across species of birds-of-paradise (S3 Table). To reduce the influence of sampling intensity on our overall behavioral analyses, we used a sliding-window analysis to evaluate similar time windows for courtship display complexity across species. Specifically, we used a sliding 50s window, chosen as the minimum duration resulting in relatively stable individual behavioral complexity scores (S2 Fig), across all clips for a given individual to identify the specific 50s period of maximal display complexity for that individual, and incorporated the resultant complexity scores for this interval in our analysis. Individual scores were then averaged to obtain species-level estimates of signal complexity. Collectively, our approach minimizes the influence that variation in recording time and clip duration has on species-level behavioral comparisons. Our results and interpretations are robust to the choice of different window sizes between 10 and 60 seconds (S3 Fig, S4 Fig, S5 Table, S6 Table).

410

415

420

Color complexity

Image collection

We collected images from 393 BOP museum specimens (S7 Table) housed at the American Museum of Natural History. Specifically, we took RAW format images of adult males from 40 BOP species under standardized conditions using a Canon 7D camera with full-spectrum quartz conversion and fitted with a Novoflex Noflexar 35 mm lens. Illumination was provided from two eyeColor arc lamps (Iwasaki: Tokyo, Japan) diffused through polytetrafluoroethylene sheets 0.5mm thick. These arc lamps are designed to simulate CIE (International Commission on Illumination) recommended daylight (D65) illumination (though they come standard with a UV-blocking coating, which we removed prior to use). Additionally, for every specimen position (see below) we used filters (Baader: Mammendorf, Germany) to take two photos, one capturing only ultraviolet light (300-400nm) and one capturing wavelengths between 400 and 700 nm.

To simulate a variety of viewing angles and increase the likelihood of capturing relevant coloration from bird specimens, we took photographs of each specimen from three viewing angles: dorsal, ventral, ventral-angled. Specifically, each specimen was photographed from above while it was flat on its belly (dorsal view), flat on its back (ventral view), and angled 45° on its back (rotating the frontal plane along the vertical-axis, while keeping the head oriented in the same direction as the previous two photographs). The angled photograph was taken to increase the likelihood of capturing some of the variation made possible by iridescent plumage.

Image processing

Ultra-violet and visible spectrum images were used to create standardized (i.e. channels were equalized and linearized [92]) multispectral image files for each specimen/position using the Image Calibration and Analysis Toolbox [93] in ImageJ [94].

a. Avian color vision

After estimating the color sensitivity of our camera/lens combination [92,93], we generated custom mapping functions to convert image colors to stimulation values corresponding to an avian visual space. Birds-of-paradise are inferred to have a VS visual system [95], and the curl-crested manucode (*Manucodia comrii*) and magnificent riflebird (*Ptiloris magnificus*) have the same amino acid sequence in spectral tuning positions 84-94 [95] as the jackdaw (*Corvus monedula*) [96], which is inferred to have a peak sensitivity of its (VS-type) SWS1 cone at 408 nm. This sensitivity is similar to that of another species with a VS visual system, the pigeon *Columbia livia* (SWS1 peak sensitivity = 404nm [97]). Consequently, we converted our full-spectrum photographs into the perceptual space of pigeons using physiological data [98] and spectral sensitivity curve functions [99,100] (implemented in the R package *pavo* [101]) and multispectral imaging software [93] in ImageJ. Additionally, we evaluated color using a visual model from a UV-sensitive passerine species (the blue-tit [102]) and found our results qualitatively unchanged. Prior to subsequent clustering (see below), we performed a median pixel blur to eliminate aberrant pixel values (owing to dust on the sensor, temporary dead-pixels, etc.).

b. Color clustering

Following conversion to avian color vision and noise filtering, we used a novel custom-written agglomerative hierarchical clustering algorithm to reduce each multispectral image down to a perceptually-relevant number of color clusters [103]. This clustering algorithm was developed from a more basic algorithm used previously [104] which did not integrate luminance, or thresholds when combining clusters. At the first step of the clustering process, each pixel is its own cluster. Each cluster is then compared to its neighboring clusters in the XY plane of the image within a given radius (1 pixel initially), and composite distances are calculated based on an equal weighting of chromatic [105] and achromatic [106] Just Noticeable Distances (using the log model). That is, chromatic and luminance JND values are divided by the chromatic JND threshold and luminance JND threshold respectively, so that they are weighted equally

based on the chosen threshold, then the Euclidean distance of these two scaled values is calculated. Following distance calculations, each cluster is combined with its nearest (in composite JND distance) neighboring cluster if the composite difference is below the threshold for both luminance and chromatic JNDs. Nodes can have multiple clustering events at each pass, e.g. if cluster A is closest to B, but B is closest to C, and all distances are below the threshold, then all three will be clustered, meaning whole strings or neighboring regions can be clustered. In practice, each cluster tends to be combined with two or three clusters on each pass. On each pass the updated mean cone catch values for each cluster are calculated, ready for the next pass. Additionally, the XY distance search radius increases with each pass, so as clusters get larger they can also combine with neighboring clusters further away which has the desirable effect of also keeping the processing for each pass relatively constant (i.e. there are fewer clusters on each pass, but each one must be compared to a larger number of neighbors). Clustering therefore takes place across $n + 2$ dimensions (n colors plus x and y space). The code used for this clustering is provided as supplementary material (S1 Code), and we have included an exemplar image illustrating the output of the clustering process (S5 Fig).

c-Measures of color complexity

Following color clustering, we quantified plumage color complexity using analogous indices to those we employed in our behavioral analysis. Namely, we quantified color richness (the number of distinct clusters and color diversity (the numbers equivalent of Shannon index) for each view (dorsal, ventral, angled) and averaged these values to obtain individual, specimen-level metrics of color complexity. In terms of color, specimens with higher richness scores have more unique colors, and species with higher diversity scores have more, evenly-distributed colors.

d-Influence of specimen age on color complexity measures

Aging can influence the coloration and appearance of some kinds of avian plumage [107,108], though such effects are often relatively small [109]. To evaluate the possibility that specimen age might influence our estimates of species' level plumage elaboration, we conducted a linear mixed-effect model with two measures of color complexity (color richness, color diversity) as the dependent variable, collection year as the independent variable, and species as a random effect. Analyzing these models revealed no significant influence of collection year on either color richness (standardized $\beta = 0.032$, 95% CI $-0.036 - 0.099$, $t = 0.930$, $p = 0.353$) or diversity (standardized $\beta = 0.033$, 95% CI $-0.012 - 0.080$, $t = 1.437$, $p = 0.151$).

Acoustic Complexity

As with display behaviors, we quantified the acoustic complexity of courtship sounds produced by analyzing field-recorded audio/video clips of 32 (80%) BOP species. In total, we analyzed sound from 176 clips from 59 individuals totaling 24670.9 seconds (≈ 411 minutes; mean clip duration = 140.18s; S8 Table). Though birds can generate sounds (both vocally and mechanically) in numerous contexts, we focused our analysis on recordings from known display sites or those matching written descriptions of courtship sound production [44].

Data collection

From each video clip used to quantify display behavior we identified a focal individual and all of the sounds it produced. Spectrograms of the audio were viewed with a frequency resolution of 43.1 Hz and time resolution of 2.31 ms, and all sounds were marked in the sound analysis software RavenPro v. 1.5[110]. Individual sounds were defined as temporally-separated sound elements. Using the robust measurements in Raven, we measured the duration, maximum and minimum frequency, bandwidth, peak frequency, and peak frequency contour of each call. We measured the disorder, lack of organized or tonal structure, in a call with aggregate entropy and average entropy measures in Raven.

Following detailed analysis of the acoustic parameters for all notes, we used a two-step semi-automated classification analysis to assign note identity. In the first step, we conducted a principal components analysis

525 using 15 summary acoustic variable (S9 Table) followed by agglomerative hierarchical clustering to assign partial note identity (i.e. note classification based on location in PCA-based sound space). In the second step, each note was given a categorical identifier depending on the combination of four qualitative variables manually scored as yes/no (frequency modulation, non-harmonic structure, impulsive, stochastic). Full note identity was achieved by merging the clustered note identity with the combined qualitative-categorization.

Measures of acoustic complexity

530 After assigning identities to all notes in our dataset, we measured acoustic richness (number of distinct note types) and acoustic diversity (Shannon index of notes) within a given time period (see *Sliding window analysis* below). As with behavior, we used the numbers equivalent of Shannon index values to facilitate more direct comparisons among samples and species.

Sliding window analysis

535 The duration and number of available courtship-specific acoustic recordings was highly variable across birds-of-paradise (S8 Table). To reduce the influence of this variation on species-level acoustic comparisons, we used a sliding-window analysis, similar to our behavioral analyses, to evaluate and compare similar time windows for acoustic display complexity across species. To identify the time period of maximal acoustic complexity for an individual in our analysis, we used a sliding 10s window, chosen as 540 the minimum duration resulting in relatively stable individual complexity scores (S6 Fig), across all clips for a given individual. Individual scores were then averaged to obtain species-level estimates of signal complexity. Relative complexity measures are robust to the choice of different window sizes between 5 and 50 seconds (S7 Fig, S8 Fig).

545

Phylogenetic analyses

Tree

550 We regenerated a phylogenetic hypothesis from a recent molecular phylogeny for Paradisaeidae [111] using the function *phylo.tracer* the R package *physketch* [112]. This ultrametric, time-scaled tree was used for all downstream comparative analyses following one modification of tree topology. Specifically, we placed *Lophorina superba* as the outgroup to *Ptiloris* to accommodate a revised taxonomic hypothesis [113].

Phylogenetic Generalized Least Squares

555 For each of the six components of courtship phenotype (color richness and diversity, behavior richness and diversity, acoustic richness and diversity), we conducted a single, multiple phylogenetic generalized least squares (mPGLS) regression evaluating the influence of the other elements of courtship phenotype and two signal-environment variables predicted to influence relative investment in separate axes of overall courtship phenotype. Specifically, we included a categorical metric of display height in all models, where each 560 species was scored as displaying on the forest floor, in the understory, or in the forest canopy. Additionally, we included the categorical metric of display proximity in all models, where each species was scored as displaying solitarily, in exploded leks, or in true leks. In each model (see S1 Table, S2 Table), we included only ‘like’ phenotype measures (e.g. including behavioral and acoustic richness, but not behavioral or acoustic diversity, when investigating the drivers of color richness). All courtship phenotype measures were 565 log-transformed prior to analyses, and analyses were performed in the R computing environment [114] using the *gls* function in the *nlme* package [115] assuming an Ornstein-Ulbeck covariance structure [116] using the *corMartins* function from the *ape* package [117].

Imputation

570 We used the *Rphylopars* package in R [118] to impute character values for taxa with missing data (e.g. species lacking behavioral/acoustic information, S4 Table). This methodology has previously been found to perform well in predicting ancestral and missing species’ values [119]. In our case, we evaluated the performance of several methods to estimate missing values assuming i) a Brownian motion model of trait

575 evolution, ii) an Ornstein-Ulbeck model [116], iii) an “early-burst” model of trait evolution [120], iv) a
Pagel’s lambda model of trait evolution [121], and v) a multivariate OU model [122]. We compared model
performance by evaluating AIC scores and determined that the OU model performed best. Consequently,
character trait values imputed using this model were used in all subsequent analysis.

580 Though data imputation can increase statistical power [123], the instances in which it might induce spurious
findings are few (especially given the relatively small proportion of our total dataset (20%) for which we
imputed values (*cf* [124,125]). In fact, bias tends to be lower when missing data are imputed rather than
omitted [119]. Regardless, to alleviate concerns that imputed values may drive subsequent findings, we also
585 conducted our phylogenetic least squares (PGLS) analyses on the limited subset of species (n=31) for which
we have complete data. In all cases, the findings were qualitatively identical to those reported in the main
text (S10 Table, S11 Table).

Data accessibility

Data for primary analyses are included in S1 Data file.

Acknowledgements

590 We thank Brett Benz, Lydia Garetano, Paul Sweet, and Joel Cracraft at the American Museum of Natural
History for their generous help. We thank the Cornell NBB “LunchBunch” for stimulating intellectual
contributions, and Rick Elliker and Matt Young for assistance with media curation. We thank Annie Lee
and Ginnie Jeng for assistance with data collection. We thank Nick Mason, Jacob Berv, Eliot Miller, Mike
595 Sheehan, Gavin Leighton, Ben Freeman, Eric Goolsby, Anastasia Dalziell, Marcelo Araya-Salas, Leo
Campagna, Sara Miller, Irby Lovette, and Mike Webster for input and advice on data collection, analysis,
and interpretation. We thank Szabolcs Kókay for the incredible paintings illustrating bird-of-paradise
display behavior in our figures. We thank Rory, Teddy, Veronica, Sandy, and David Ligon for emotional
support.

600

Literature Cited

1. Schluter D. *The Ecology of Adaptive Radiation*. New York: Oxford University Press; 2000.
2. Seehausen O, Terai Y, Magalhaes IS, Carleton KL, Mrosso HDJ, Miyagi R, et al. Speciation through sensory drive in cichlid fish. *Nature*. 2008;455: 620–6. doi:10.1038/nature07285
3. Wagner CE, Harmon LJ, Seehausen O. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature*. Nature Publishing Group; 2012;487: 366–9. doi:10.1038/nature11144
4. Hofmann CM, O’Quin KE, Marshall NJ, Cronin TW, Seehausen O, Carleton KL. The eyes have it: regulatory and structural changes both underlie cichlid visual pigment diversity. *PLoS Biol*. 2009;7: e1000266. doi:10.1371/journal.pbio.1000266
5. Carlson BA, Hasan SM, Hollmann M, Miller DB, Harmon LJ, Arnegard ME. Brain evolution triggers increased diversification of electric fishes. *Science* (80-). 2011;332: 583–586. doi:10.1126/science.1201524
6. Price JJ, Friedman NR, Omland KE. Song and plumage evolution in the New World orioles (*Icterus*) show similar lability and convergence in patterns. *Evolution* (N Y). 2007;61: 850–863. doi:10.1111/j.1558-5646.2007.00082.x
7. Mendelson TC, Martin MD, Flaxman SM. Mutation-order divergence by sexual selection: Diversification of sexual signals in similar environments as a first step in speciation. *Ecol Lett*. 2014;17: 1053–1066. doi:10.1111/ele.12313
8. Fisher RA. *The Genetical Theory of Natural Selection*. Oxford University Press; 1930. doi:10.1038/158453a0
9. Lande R. Models of speciation by sexual selection on polygenic traits. *Proc Natl Acad Sci U S A*. 1981;78: 3721–3725. doi:10.1073/pnas.78.6.3721
10. Pomiankowski A, Iwasa Y. Runaway ornament diversity caused by Fisherian sexual selection. *Proc Natl Acad Sci U S A*. 1998;95: 5106–5111. doi:10.1073/pnas.95.9.5106
11. Kirkpatrick M. Sexual selection and the evolution of female choice. *Evolution* (N Y). 1982;36: 1–12. doi:10.2307/2407961
12. Møller AP, Pomiankowski A. Why have birds got multiple sexual ornaments? *Behav Ecol Sociobiol*. 1993;32: 167–176.
13. Iwasa Y, Pomiankowski A. The Evolution of Mate Preferences for Multiple Sexual Ornaments. *Evolution* (N Y). 1994;48: 853–867.
14. Hebets EA, Papaj DR. Complex signal function: developing a framework of testable hypotheses. *Behav Ecol Sociobiol*. 2005;57: 197–214. doi:10.1007/s00265-004-0865-7
15. Rubi TL, Stephens DW. Why Complex Signals Matter, Sometimes. In: Bee MA, Miller CT, editors. *Psychological Mechanisms in Animal Communication*. Vol. 5. Springer International Publishing; 2016. pp. 119–136. doi:10.1007/978-3-319-48690-1_1
16. Johnstone RA. Multiple displays in animal communication: “backup signals” and “multiple messages.” *Philos Trans R Soc B*. 1996;351: 329–338.
17. Pomiankowski A, Iwasa Y. Evolution of Multiple Sexual Preferences by Fisher’s Runaway Process of Sexual Selection. *Proc R Soc B*. 1993;253: 173–181. doi:10.1098/rspb.1993.0099
18. Moller AP, Pomiankowski A. Why have birds got multiple sexual ornaments? *Behav Ecol Sociobiol*. 1993;32: 167–176.
19. Dale J, Dey C, Delhey K, Kempnaers B, Valcu M. The effects of life-history and social selection on male and female plumage coloration. *Nature*. Nature Publishing Group; 2015;527: 367–370. doi:10.1038/nature15509
20. Arnegard ME, McIntyre PB, Harmon LJ, Zelditch ML, Crampton WGR, Davis JK, et al. Sexual Signal Evolution Outpaces Ecological Divergence during Electric Fish Species Radiation. *Am Nat*. 2010;176: 335–356. doi:10.1086/655221
21. Tobias JA, Cornwallis CK, Derryberry EP, Claramunt S, Brumfield RT, Seddon N. Species coexistence and the dynamics of phenotypic evolution in adaptive radiation. *Nature*. Nature Publishing Group; 2014;506: 359–363. doi:10.1038/nature12874

22. Johnson KP. The evolution of bill coloration and plumage dimorphism supports the transference hypothesis in dabbling ducks. *Behav Ecol.* 1999;10: 63–67. doi:10.1093/beheco/10.1.63
- 655 23. Gilliard ET. Bower ornamentation versus plumage characters in bower-birds. *Auk.* 1956;73: 450–451.
24. Kusmierski R, Borgia G, Uy A, Crozier RH. Labile evolution of display traits in bowerbirds indicates reduced effects of phylogenetic constraint. *Proc R Soc B.* 1997;264: 307–13. doi:10.1098/rspb.1997.0044
- 660 25. Badyaev A V, Hill GE, Weckworth B V. Species Divergence in Sexually Selected Traits: Increase in Song Elaboration is Related to Decrease in Plumage Ornamentation in Finches. *Evolution (N Y).* 2002;56: 412–419.
26. Galván I. The importance of white on black: unmelanized plumage proportion predicts display complexity in birds. *Behav Ecol Sociobiol.* 2008;63: 303–311. doi:10.1007/s00265-008-0662-9
- 665 27. Wagner WE, Beckers OM, Tolle AE, Basolo AL. Tradeoffs limit the evolution of male traits that are attractive to females. *Proc R Soc B.* 2012;279: 2899–2906. doi:10.1098/rspb.2012.0275
28. Ossip-Klein AG, Fuentes JA, Hews DK, Martins EP. Information content is more important than sensory system or physical distance in guiding the long-term evolutionary relationships between signaling modalities in *Sceloporus* lizards. *Behav Ecol Sociobiol.* 2013;67: 1513–1522. doi:10.1007/s00265-013-1535-4
- 670 29. Darwin C. *Secondary Sexual Characters of Birds. The Descent of Man & Selection in Relation to Sex.* 1st ed. London: John Murray; 1871. pp. 38–98.
30. Wallace AR. *Colours and ornaments characteristic of sex. Darwinism: An Exposition of the Theory of Natural Selection with Some of Its Applications.* 1889. pp. 268–300.
- 675 31. Mason NA, Shultz AJ, Burns KJ. Elaborate visual and acoustic signals evolve independently in a large , phenotypically diverse radiation of songbirds. *Proc R Soc B.* 2014;281: 1–10. doi:10.1098/rspb.2014.0967
32. Ornelas JF, González C, Espinosa De Los Monteros a. Uncorrelated evolution between vocal and plumage coloration traits in the trogons: A comparative study. *J Evol Biol.* 2009;22: 471–484. doi:10.1111/j.1420-9101.2008.01679.x
- 680 33. Matysioková B, Friedman N, Turčoková L, Remeš V. The evolution of feather coloration and song in Old World orioles (genus *Oriolus*). *J Avian Biol.* 2017;48: 1015–1024. doi:10.1111/oik.02986
34. Gomes ACR, Funghi C, Soma M, Sorenson MD, Cardoso GC. Multi-modal signalling in estrildid finches: song, dance and colour are associated with different ecological and life history traits. *J Evol Biol.* 2017;30: 1–11. doi:10.1111/jeb.13102
- 685 35. Soma M, Garamszegi LZ. Evolution of courtship display in Estrildid finches: dance in relation to female song and plumage ornamentation. *Front Ecol Evol.* 2015;3: 1–11. doi:10.3389/fevo.2015.00004
- 690 36. Wiens JJ. Decoupled evolution of display morphology and display behaviour in phrynosomatid lizards. *Biol J Linn Soc.* 2000;70: 597–612. doi:10.1006/hijl.
37. Doorn GS Van, Weissing FJ. The Evolution of Female Preferences for Multiple Indicators of Quality The Evolution of Female Preferences for Multiple Indicators of Quality. *Am Nat.* 2004;164: 173–186.
- 695 38. Shutler D, Weatherhead PJ. *Targets of Sexual Selection: Song and Plumage of Wood Warblers* Author. *Evolution (N Y).* 1990;44: 1967–1977.
39. Webb WH, Brunton DH, Aguirre JD, Thomas DB, Valcu M, Dale J. Female Song Occurs in Songbirds with More Elaborate Female Coloration and Reduced Sexual Dichromatism. *Front Ecol Evol.* 2016;4: 1–8. doi:10.3389/fevo.2016.00022
- 700 40. Badyaev A V. Integration and Modularity in the Evolution of Sexual Ornaments. In: Pigliucci M, Preston K, editors. *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes.* Oxford University Press; 2004. pp. 50–79.
41. Peiman KS, Robinson BW. Comparative Analyses of Phenotypic Trait Covariation within and

- among Populations. *Am Nat.* 2017;190: 451–468. doi:10.1086/693482
- 705 42. Stafstrom JA, Hebets EA. Female mate choice for multimodal courtship and the importance of the signaling background for selection on male ornamentation. *Curr Zool.* 2013;59: 200–209. doi:10.1093/59.2.200
43. Kasurak A V., Zielinski BS, Higgs DM. Reproductive status influences multisensory integration responses in female round gobies, *Neogobius melanostomus*. *Anim Behav.* Elsevier Ltd; 2012;83: 1179–1185. doi:10.1016/j.anbehav.2012.02.008
- 710 44. Frith CB, Beehler BM. *The birds of paradise: paradisaeidae*. Oxford: Oxford University Press; 1998.
45. Laman T, Scholes III E. *Birds of paradise: revealing the world's most extraordinary birds*. Washington, DC: National Geographic; 2014.
- 715 46. Irestedt M, Jönsson K a, Fjeldså J, Christidis L, Ericson PGP. An unexpectedly long history of sexual selection in birds-of-paradise. *BMC Evol Biol.* 2009;9: 235. doi:10.1186/1471-2148-9-235
47. Ligon JD. *The Evolution of Avian Breeding Systems*. New York: Oxford University Press; 1999.
48. Wilts BD, Michielsen K, De Raedt H, Stavenga DG. Sparkling feather reflections of a bird-of-paradise explained by finite-difference time-domain modeling. *Proc Natl Acad Sci.* 2014;111: 4363–4368. doi:10.1073/pnas.1323611111
- 720 49. Stavenga DG, Leertouwer HL, Marshall NJ, Osorio D. Dramatic colour changes in a bird of paradise caused by uniquely structured breast feather barbules. *Proc R Soc B.* 2011;278: 2098–2104. doi:10.1098/rspb.2010.2293
- 725 50. McCoy DE, Feo T, Harvey TA, Prum RO. Structural absorption by barbule microstructures of super black bird of paradise feathers. *Nat Commun.* Springer US; 2018;9: 1. doi:10.1038/s41467-017-02088-w
51. Lecroy M. *The Genus Paradisaea-Display and Evolution*. *Am Museum Novit.* 1981; 1–52.
52. Scholes Iii E. Evolution of the courtship phenotype in the bird of paradise genus *Parotia* (Aves: Paradisaeidae): homology, phylogeny, and modularity. *Biol J Linn Soc.* 2008;94: 491–504. doi:10.1111/j.1095-8312.2008.01012.x
- 730 53. Laman, T. & Scholes E. *Birds of paradise: revealing the world's most extraordinary birds*. Washington, DC: National Geographic Books; 2012.
54. Endler JA. Signals, Signal Conditions, and the Direction of Evolution. *Am Nat.* 1992;139: S125–S153.
- 735 55. Morton ES. Ecological Sources of Selection on Avian Sounds. *Am Nat.* 1975;109: 17–34.
56. Wagner GP, Schwenk K. Evolutionarily Stable Configurations: Functional Integration and the Evolution of Phenotypic Stability. In: Hecht MK, Macintyre RJ, Clegg MT, editors. *Evolutionary Biology Vol 31*. New York: Springer Science+Business Media; 2000. pp. 155–218.
57. Mayr E. *The Emergence of Evolutionary Novelties. Evolution and the Diversity of Life*. Cambridge, MA: Harvard University Press; 1976. pp. 88–113.
- 740 58. Darwin C. *The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 6th ed. London: Murray; 1872.
59. Hebets EA, Barron AB, Balakrishnan CN, Hauber ME, Mason PH, Hoke KL. A systems approach to animal communication. *Proc R Soc B.* 2016;283: 20152889. doi:10.1098/rspb.2015.2889
- 745 60. de Visser JAGM, Hermisson J, Wagner GP, Meyers LA, Bagheri-Chaichian H, Blanchard JL, et al. Perspective: Evolution and Detection of Genetic Robustness. *Evolution (N Y)*. 2003;57: 1959. doi:10.1554/02-750R
61. Wagner A. The role of robustness in phenotypic adaptation and innovation. *Proc R Soc B.* 2012;279: 1249–1258. doi:10.1098/rspb.2011.2293
- 750 62. Wagner A. Distributed robustness versus redundancy as causes of mutational robustness. *BioEssays.* 2005;27: 176–188. doi:10.1002/bies.20170
63. Ibáñez-Marcelo E, Alarcón T. The topology of robustness and evolvability in evolutionary systems with genotype-phenotype map. *J Theor Biol.* Elsevier; 2014;356: 144–162. doi:10.1016/j.jtbi.2014.04.014

- 755 64. Wagner A. Robustness, evolvability, and neutrality. *FEBS Lett.* 2005;579: 1772–1778. doi:10.1016/j.febslet.2005.01.063
65. Toll-Riera M, San Millan A, Wagner A, MacLean RC. The Genomic Basis of Evolutionary Innovation in *Pseudomonas aeruginosa*. *PLoS Genet.* 2016;12: 1–21. doi:10.1371/journal.pgen.1006005
- 760 66. Raman K, Wagner A. Evolvability and robustness in a complex signalling circuit. *Mol Biosyst.* 2011;7: 1081. doi:10.1039/c0mb00165a
67. Payne JL, Wagner A. The Robustness and Evolvability of Transcription Factor Binding Sites. *Science (80-)*. 2014;343: 875–877. doi:10.1126/science.1249046
68. Ay N, Flack J, Krakauer DC. Robustness and complexity co-constructed in multimodal signalling networks. *Philos Trans R Soc B.* 2007;362: 441–447. doi:10.1098/rstb.2006.1971
- 765 69. Schluter D. Evidence for Ecological Speciation and Its Alternative. *Science (80-)*. 2009;323: 737–741. doi:10.1126/science.1160006
70. Nosil P, Flaxman SM. Conditions for mutation-order speciation. *Proc R Soc B.* 2011;278: 399–407. doi:10.1098/rspb.2010.1215
- 770 71. Prum RO. Aesthetic evolution by mate choice: Darwin’s really dangerous idea. *Philos Trans R Soc B Biol Sci.* 2012;367: 2253–2265. doi:10.1098/rstb.2011.0285
72. Arnold SJ, Houck LD. Can the Fisher-Lande Process Account for Birds of Paradise and Other Sexual Radiations? *Am Nat.* 2016;187: 717–735. doi:10.1086/686258
73. Diamond J. Biology of Birds of Paradise and Bowerbirds. *Annu Rev Ecol Syst.* 1986;17: 17–37.
- 775 74. Miles MC, Fuxjager MJ. Synergistic selection regimens drive the evolution of display complexity in birds of paradise. *J Anim Ecol.* 2018; 1–11. doi:10.1111/1365-2656.12824
75. Endler JA. Some general comments on the evolution and design of animal communication systems. *Philos Trans R Soc B.* 1993;340: 215–225. doi:10.1098/rstb.1993.0060
76. Candolin U. The use of multiple cues in mate choice. *Biol Rev.* 2003;78: 575–95.
- 780 77. Endler JA, Basolo AL. Sensory ecology, receiver biases and sexual selection. *Trends Ecol Evol.* 1998;13: 415–420.
78. Akre KL, Johnsen S. Psychophysics and the evolution of behavior. *Trends Ecol Evol.* 2014; 1–10. doi:10.1016/j.tree.2014.03.007
79. Guilford T, Dawkins MS. Receiver psychology and the evolution of animal signals. *Anim Behav.* 1991;42: 1–14.
- 785 80. Rowe C. Receiver psychology and the evolution of multicomponent signals. *Anim Behav.* 1999;58: 921–931. doi:10.1006/anbe.1999.1242
81. Hillis JM, Ernst MO, Banks MS, Landy MS. Combining sensory information: mandatory fusion within, but not between, senses. *Science.* 2002;298: 1627–1630. doi:10.1126/science.1075396
- 790 82. Rubi TL, Stephens DW. Does multimodality per se improve receiver performance? An explicit comparison of multimodal versus unimodal complex signals in a learned signal following task. *Behav Ecol Sociobiol.* 2016;70: 409–416. doi:10.1007/s00265-016-2061-y
83. Hill GE. Sexiness, Individual Condition, and Species Identity: The Information Signaled by Ornaments and Assessed by Choosing Females. *Evol Biol.* Springer US; 2015;42: 251–259. doi:10.1007/s11692-015-9331-x
- 795 84. Pruett-Jones SG, Pruett-Jones MA. Sexual Selection Through Female Choice in Lawes’ Parotia, A Lek-Mating Bird of Paradise. *Evolution (N Y).* 1990;44: 486–501.
85. Prum RO. Phylogenetic analysis of the evolution of display behavior in the Neotropical manakins (Aves; Pipridae).pdf. *Ethology.* 1990;84: 202–231.
- 800 86. Martins EP. A Comparative Study of the Evolution of Sceloporus Push-Up Displays. *Am Nat.* 1993;142: 994–1018.
87. Stearns SC. Trade-Offs in Life-History Evolution. *Funct Ecol.* 1989;3: 259–268.
88. Hänninen L, Pastell M. CowLog: Open source software for coding behaviors from digital video. *Behav Res Methods.* 2009;41: 472–476.
- 805 89. Shannon CE. A Mathematical Theory of Communication. *Bell Syst Tech J.* 1948;27: 379–423,

- 623–656.
90. Jost L. Partitioning diversity into independent alpha and beta components. *Ecology*. 2007;88: 2427–2439.
91. MacArthur RH. Patterns of species diversity. *Biol Rev*. 1965;40: 510–533.
- 810 92. Stevens M, Párraga CA, Cuthill IC, Partridge JC, Troscianko TOMS, Psychology E, et al. Using digital photography to study animal coloration. *Biol J Linn Soc*. 2007;90: 211–237.
93. Troscianko J, Stevens M. Image calibration and analysis toolbox - a free software suite for objectively measuring reflectance, colour and pattern. *Methods Ecol Evol*. 2015;6: 1320–1331. doi:10.1111/2041-210X.12439
- 815 94. Schneider C a, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. Nature Publishing Group; 2012;9: 671–675. doi:10.1038/nmeth.2089
95. Ödeen A, Håstad O, Alström P. Evolution of ultraviolet vision in the largest avian radiation - the passerines. *BMC Evol Biol*. BioMed Central Ltd; 2011;11: 313. doi:10.1186/1471-2148-11-313
- 820 96. Odeen A, Hastad O. Complex distribution of avian color vision systems revealed by sequencing the SWS1 opsin from total DNA. *Mol Biol Evol*. 2003;20: 855–61. doi:10.1093/molbev/msg108
97. Hart NS, Hunt DM. Avian Visual Pigments: Characteristics , Spectral Tuning, and Evolution. *Am Nat*. 2007;169: S7–S26.
98. Bowmaker J, Heath L, Wilkie S, Hunt D. Visual Pigments and Oil Droplets from Six Classes of Photoreceptor in the Retinas of Birds. *Vision Res*. 1997;37: 2183–2194.
- 825 99. Govardovskii VI, Fyhrquist N, Reuter T, Kuzmin DG, Donner K. In search of the visual pigment template. *Vis Neurosci*. 2000;17: 509–28.
100. Hart NS, Vorobyev M. Modelling oil droplet absorption spectra and spectral sensitivities of bird cone photoreceptors. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol*. 2005;191: 381–92. doi:10.1007/s00359-004-0595-3
- 830 101. Maia R, Eliason CM, Bitton P-P, Doucet SM, Shawkey MD. pavo : an R package for the analysis, visualization and organization of spectral data. Tatem A, editor. *Methods Ecol Evol*. 2013; n/a-n/a. doi:10.1111/2041-210X.12069
102. Hart NS, Partridge JC, Cuthill IC, Bennett AT. Visual pigments, oil droplets, ocular media and cone photoreceptor distribution in two species of passerine bird: the blue tit (*Parus caeruleus* L.) and the blackbird (*Turdus merula* L.). *J Comp Physiol A*. 2000;186: 375–87.
- 835 103. van den Berg CP, Cheney KL, Endler JA, Marshall JN, Troscianko J. Image Smoothing and Segmentation using Colour and Luminance Discrimination Thresholds: JND Clustering and Weber Smart Blur. (In prep).
104. Troscianko J, Claire JW, Stevens M. Nest covering in plovers : How modifying the visual environment influences egg camouflage. *Ecol Evol*. 2016; 7536–7545. doi:10.1002/ece3.2494
- 840 105. Vorobyev M, Osorio D. Receptor noise as a determinant of colour thresholds. *Proc R Soc B*. 1998;265: 351–358. doi:10.1098/rspb.1998.0302
106. Siddiqi A, Cronin TW, Loew ER, Vorobyev M, Summers K. Interspecific and intraspecific views of color signals in the strawberry poison frog *Dendrobates pumilio*. *J Exp Biol*. 2004;207: 2471–2485. doi:10.1242/jeb.01047
- 845 107. McNett GD, Marchetti K. Ultraviolet degradation in carotenoid patches: live versus museum specimens of wood warblers (Parulidae). *Auk*. 2005;122: 793–802. doi:10.2307/4090586
108. Doucet SM, Hill GE. Do museum specimens accurately represent wild birds? A case study of carotenoid, melanin, and structural colours in long-tailed manakins *Chiroxiphia linearis*. *J Avian Biol*. 2009;40: 146–156. doi:10.1111/j.1600-048X.2009.03763.x
- 850 109. Armenta JK, Dunn, PO, Whittingham LA. Effects of Specimen Age on Plumage Color. *Auk*. 2008;125: 803–808. doi:10.1525/auk.2008.07006
110. Bioacoustics Research Program. Raven Pro: Interactive Sound Analysis Software (Version 1.5). Ithaca, NY: The Cornell Lab of Ornithology. Available from <http://www.birds.cornell.edu/raven.>; 2014.
- 855 111. Irestedt M, Jönsson K a, Fjeldså J, Christidis L, Ericson PGP. An unexpectedly long history of

- sexual selection in birds-of-paradise. *BMC Evol Biol.* 2009;9: 235. doi:10.1186/1471-2148-9-235
112. Revell LJ. physketch: Drawing phylogenetic objects free-hand. R Packag. 2017; <http://github.com/liamrevell/physketch>.
- 860 113. Irestedt M, Batalha-Filho H, Fls PGPE, Christidis L, Schodde R. Phylogeny, biogeography and taxonomic consequences in a bird-of-paradise species complex, *Lophorina – Ptiloris* (Aves: Paradisaeidae). *Zool J Linn Soc.* 2017;181: 439–470.
114. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. 2015; <http://www.R-project.org/>.
- 865 115. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. R Packag version 31-1311. 2018; <https://CRAN.R-project.org/package=nlme>.
116. Martins EP, Hansen TF. Phylogenies and the Comparative Method: A General Approach to Incorporating Phylogenetic Information into the Analysis of Interspecific Data. *Am Nat.* 1997;149: 646–667.
- 870 117. Paradis E, Strimmer K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics.* 2004;20: 289–290.
118. Goolsby EW, Bruggeman J, Ané C. Rphylopars : fast multivariate phylogenetic comparative methods for missing data and within-species variation. *Methods Ecol Evol.* 2017;8: 22–27. doi:10.1111/2041-210X.12612
- 875 119. Penone C, Davidson AD, Shoemaker KT, Marco M Di, Brooks TM, Young BE, et al. Imputation of missing data in life-history trait datasets: which approach performs the best? *Methods Ecol Evol.* 2014;5: 961–970. doi:10.1111/2041-210X.12232
120. Harmon LJ, Losos JB, Jonathan Davies T, Gillespie RG, Gittleman JL, Bryan Jennings W, et al. Early Bursts of Body Size and Shape Evolution Are Rare in Comparative Data: Early Bursts of Evolution Are Rare. *Evolution (N Y).* 2010;64: 2385–2396. doi:10.1111/j.1558-5646.2010.01025.x
- 880 121. Pagel M. Inferring evolutionary processes from phylogenies. *Zool Scr.* 1997;26: 331–348. doi:10.1111/j.1463-6409.1997.tb00423.x
122. Bartoszek K, Pienaar J, Mostad P, Andersson S, Hansen TF. A phylogenetic comparative method for studying multivariate adaptation. *J Theor Biol. Elsevier;* 2012;314: 204–215. doi:10.1016/j.jtbi.2012.08.005
- 885 123. Nakagawa S, Freckleton RP. Missing inaction : the dangers of ignoring missing data . *Trends Ecol Evol Missing inaction : the dangers of ignoring missing data.* 2015; 2–6. doi:10.1016/j.tree.2008.06.014
- 890 124. Pollux BJA, Meredith RW, Springer MS, Garland T, Reznick DN. The evolution of the placenta drives a shift in sexual selection in livebearing fish. *Nature. Nature Publishing Group;* 2014;513: 233–236. doi:10.1038/nature13451
125. Uyeda JC, Pennell MW, Miller ET, Maia R, McClain CR. The Evolution of Energetic Scaling across the Vertebrate Tree of Life. *Am Nat.* 2017;190: 185–199. doi:10.1086/692326
- 895

SUPPORTING INFORMATION LEGEND

900 **S1 Video : Bird-of-paradise behavioral scoring demonstration. In this video, a male western parotia**
***Parotia sefilata* performs a species-typical courtship dance for females perching above.** This video
demonstrates the concordance between a sub-sample of our scored behaviors and the actual performance
of the bird in real-time. The behaviors represent body-position moving (BP1), changing direction while
moving (BP2), shape-shifting (SS1), bowing (O3), ornamental head plumage accentuation by moving those
feathers (OPMH), ornamental flank plumage accentuation by moving those feathers (OPMF), ornamental
905 head plumage by moving the head (OPAH), and ornamental flank plumage accentuation by moving the
torso (OPAC1). Users who cannot download the video can also view it here:
<https://youtu.be/MdqUO1RtbP0>

910 **S1 Code.** ImageJ plugin (java) for Hierarchical clustering using chromatic and achromatic JNDs.

S1 Data. Species-specific courtship phenotype estimates.

915 **S1 Table.** Multiple phylogenetic least-squares (mPGLS) analyses of communication-relevant influences
on three axes of courtship phenotype diversity. Categorical comparisons of display site are made with
respect to ground-displaying birds, and breeding system comparisons are made with respect to solitarily-
displaying birds.

920 **S2 Table.** Multiple phylogenetic least-squares (mPGLS) analyses of communication-relevant influences
on three axes of courtship phenotype richness. Categorical comparisons of display site are made with
respect to ground-displaying birds, and breeding system comparisons are made with respect to solitarily-
displaying birds.

S3 Table: Species sampled for courtship behavior, including the number of individuals watched.

925 **S4 Table.** Ethogram describing behavioral subunits scored while observing courtship display behavior of
birds-of-paradise.

930 **S5 Table.** Multiple phylogenetic least-squares (mPGLS) analyses of communication-relevant influences
on three axes of courtship phenotype diversity conducted using behavioral complexity metrics from a 10s
and 60s time-window. For comparison, the analyses presented in the main text focus on behavioral
complexity estimated from a 50s time-window.

935 **S6 Table.** Multiple phylogenetic least-squares (mPGLS) analyses of communication-relevant influences
on three axes of courtship phenotype richness conducted using behavioral complexity metrics from 10s
and 60s time-windows. The analyses presented in the main text focus on behavioral complexity estimated
from a 50s time-window.

940 **S7 Table.** Summary of specimens located at the American Museum of Natural History used to
quantify color complexity in the birds-of-paradise.

S8 Table. Species sampled for acoustic courtship complexity, including the number of individuals
watched.

945 **S9 Table.** Partial summary (PC1-PC3) of principal components analysis of 5739 notes produced by 32
BOP species. PC loadings for PC1-PC3 were used to plot notes in 3-dimensional PCA-space prior to
agglomerative hierarchical clustering based on Euclidean distances to categorize notes.

950 **S10 Table.** Multiple phylogenetic least-squares (mPGLS) analyses of communication-relevant influences on three axes of courtship phenotype diversity conducted only on species without imputed species level values.

955 **S11 Table.** Multiple phylogenetic least-squares (mPGLS) analyses of communication-relevant influences on three axes of courtship phenotype richness conducted only on species without imputed species level values.

960 **S1 Fig. There is no evidence for correlated evolution between color and behavioral diversity among birds-of-paradise.** Multiple phylogenetic least squares regression (mPGLS) reveals no significant relationship between behavioral and color diversity when controlling for acoustic diversity, display height, and display proximity. This plot is a phylo-signal-space plot where species ornamentation values plotted with colored circles corresponding display environment and mating system and are connected based on their phylogenetic relationships. Species' locations represent tip values for log transformed behavioral and color diversity. Underlying data for S1 Fig can be found in S1 Data.

965 **S2 Fig.** Accumulation of unique behaviors plateaus by time windows of approximately 50 seconds for most species. Note, these are unique behaviors per individual (not per clip). This distinction is important because some individuals were recorded in several clips, but in the longer clips they might not be doing much behaviorally (leading to the initially surprising drop in unique behaviors at certain longer window sizes).

970 **S3 Fig.** Pairwise comparisons of behavioral richness (number of unique behaviors) estimates for windows between 10 and 60 seconds in duration. Within plots, each point represents a species in the family Paradisaeidae, with species-specific values obtained from rphylopars reconstructions incorporating intra- and interspecific variation. Best-fit lines in lower plots, as well as F and P values presented in corresponding upper-diagonal squares, come from PGLS (phylogenetic generalized least squares) analysis assuming Ornstein–Uhlenbeck error structure. Results are qualitatively identical assuming different correlation structures (e.g. Pagel, Brownian). Underlying data for S3 Fig can be found in S2 Data.

975 **S4 Fig.** Pairwise comparisons of behavioral diversity (Shannon indices of behaviors) estimates for windows between 10 and 60 seconds in duration. Within plots, each point represents a species in the family Paradisaeidae, with species-specific values obtained from rphylopars reconstructions incorporating intra- and interspecific variation. Best-fit lines in lower plots, as well as F and P values presented in corresponding upper-diagonal squares, come from PGLS (phylogenetic generalized least squares) analysis assuming Ornstein–Uhlenbeck error structure. Results are qualitatively identical assuming different correlation structures (e.g. Pagel, Brownian). Underlying data for S4 Fig can be found in S3 Data.

985 **S5 Fig. Dorsal view of raw (left side) and clustered (right side) images taken of a Wilson's bird-of-paradise.** Following clustering based on chromatic and achromatic thresholds (see Methods), every pixel in every image is assigned to a categorical color identity. The total number of colors in an image provides a measure of richness, and the numbers equivalent of the Shannon diversity of the colors, taking into account the relative area covered by each class of colors, provides a measure of color diversity. Individuals with higher richness scores have more colors, and individuals with more colors, more evenly distributed in-terms of their relative areas, have higher diversity scores.

990 **S6 Fig.** Accumulation of unique sounds plateaus at time windows of approximately 10 s for most species.

995 **S7 Fig.** Pairwise comparisons of acoustic richness (number of unique note types) estimates for windows between 5 and 50 seconds in duration. Within plots, each point represents a species in the family Paradisaeidae, with species-specific values obtained from rphylopars reconstructions incorporating intra- and interspecific variation. Best-fit lines in lower plots, as well as F and P values presented in corresponding upper-diagonal squares, come from PGLS (phylogenetic generalized least squares) analysis assuming Ornstein–Uhlenbeck error structure. Results are

1000 qualitatively identical assuming different correlation structures (e.g. Pagel, Brownian). Underlying data for S7 Fig can be found in S4 Data.

1005 **S8 Fig.** Pairwise comparisons of acoustic diversity (log transformed Shannon indices of note complexity) estimates for windows between 5 and 50 seconds in duration. Within plots, each point represents a species in the family Paradisaeidae, with species-specific values obtained from rphylopars reconstructions incorporating intra- and interspecific variation. Best-fit lines in lower plots, as well as F and P values presented in corresponding upper-diagonal squares, come from PGLS (phylogenetic generalized least squares) analysis assuming Ornstein–Uhlenbeck error structure. Results are qualitatively identical assuming different correlation structures (e.g. Pagel, Brownian).
1010 Underlying data for S8 Fig can be found in S5 Data.