MODULATION OF AVIAN IRIDESCENCE VIA MELANOGENESIS

Soraia Barbosa[1,2,†,*], Carina Bittner[3†], Roberto Arbore[1,2,4†], Pedro M. Araújo[1,2,5†], Paulo Pereira[1,2], Pedro Andrade[1,2], Csaba Fekete[6], Rita Afonso[1,2,7], Sandra Afonso[1,2], Maria F. Amorim[1,2,7], Cristiana I. Marques[1,2,7], Michaël P. J. Nicolaï[8], João Carita[9], Jindřich Brejcha[10], Ricardo J. Lopes[11], Joel M. Alves[1,2,12], Fernando Cruz[13,14], Jèssica Gómez-Garrido[13,14], Cristina Zamarreño[13,15], Marta Gut[13,14], Tyler Alioto[13,14], Muhammad Akhtar Ali[16,17], Andrea Hilpert[18], Anita Hech[18], Erdmann Spiecker[19,20], Benjamin Apeleo Zubiri[19,20], Shosuke Ito[21], Kazumasa Wakamatsu[21], Leif Andersson[16,22], Joseph C. Corbo[4], Nicolas Vogel[3,‡,*], Miguel Carneiro[1,2,‡,*]

[1]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Universidade do Porto, Vairão, Portugal
[2]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Vairão, Portugal
[3]Institute of Particle Technology, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany
[4]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA
[5]CFE – Centre for Functional Ecology: Science for People & Planet, Marine Resources, Conservation and Technology, Department of Life Sciences, Coimbra, Portugal
[6]Laboratory of Integrative Neuroendocrinology, HUN-REN Institute of Experimental Medicine, Budapest, Hungary
[7]Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal
[8]Evolution and Optics of Nanostructures Group, Biology Department, Ghent University, Ghent, Belgium.
[9]Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Lisboa, Portugal
[10]Department of Philosophy and History of Science, Faculty of Science, Charles University in Prague, Praha, Czech Republic
[11]CE3C, Center for Ecology, Evolution and Environmental Change & CHANGE, Global Change and Sustainability Institute, Departamento de Biologia, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal
[12]Palaeogenomics and Bio-Archaeology Research Network, School of Archaeology, University of Oxford, Oxford, OX1 3QY, United Kingdom
[13]Centro Nacional de Análisis Genómico (CNAG), Baldiri Reixac 4, 08028 Barcelona, Spain
[14]Universitat de Barcelona (UB), Barcelona, Spain
[15]Universitat Autónoma de Barcelona (UAB), 08193 Bellaterra (Cerdanyola del Vallès), Barcelona, Spain
[16]Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden
[17]School of Biological Sciences, University of the Punjab, Lahore, Pakistan
[18]Friedrich-Alexander-Universität Erlangen-Nürnberg, Institut für Mikroskopische Anatomie und Molekulare Bildgebung
[19]Friedrich-Alexander-Universität Erlangen-Nürnberg, Department of Materials Science & Engineering, Institute of Micro- and Nanostructure Research (IMN)
[20]Center for Nanoanalysis and Electron Microscopy (CENEM), IZNF, 91058 Erlangen, Germany
[21]Institute for Melanin Chemistry, Fujita Health University, Toyoake, Aichi, Japan
[22]Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA
[†]Co-first authors
[‡]Co-senior authors
***Correspondence**: miguel.carneiro@cibio.up.pt (M.C.), nicolas.vogel@fau.de (N.V.); soraiabarbosa@cibio.up.pt (S.B.)

Abstract

The iridescent colors of birds originate in nanoscale feather structures that interact with light. Although the physical principles governing avian iridescence are well-established, the molecular mechanisms assembling these nanostructures into photonic materials remain unknown. Here, we investigate the genetic and developmental basis of iridescence by analyzing mutations from peafowl domestication and color variation in wild birds. We show that melanogenesis governs diverse geometric aspects of feather photonic nanostructures, enabling gains, shifts, and losses in iridescence. Mutations altering melanin composition collapse multilayered photonic systems and yield non-iridescent tissues, while alterations to melanosome abundance, elongation, or deposition timing generate multilayer architectures of variable periodicity and striking color differences. Notably, transitions from non-iridescent to iridescent plumage can arise from single-nucleotide mutations in melanogenic genes, revealing that minimal genetic change can prompt feathers to organize photonic structures. These results demonstrate that the nanoscale order underlying iridescence is developmentally plastic, emerging from physicochemical self-assembly responsive to the biochemical environment rather than from genetically encoded spatial cues. Single-cell transcriptomes from wild species further uncovered extensive melanocyte-centered regulatory rewiring associated with iridescence. Our findings identify melanogenesis as a key pathway for evolving structural coloration and illustrate how genetic programs harness self-organizing processes to generate biological diversity.

Introduction

From butterfly wings to bird feathers, animal coloration arises through two primary mechanisms: pigmentary and structural coloration (*1–4*). In birds, pigments generate hues ranging from blacks and browns to bright reds and yellows (*5*), and their genetic and biochemical underpinnings are increasingly well understood (*6–9*). However, some of the most visually striking feather effects—glossy blacks, bright blues, and iridescence—result from structural coloration (*10–12*). This form of coloration originates from the interaction of light with highly organized nanoscale features within feathers (*10*, *13*), and evolves more rapidly than pigment-based colors (*14–19*). Notable examples include the elaborate plumage of hummingbirds and peacocks, whose eyespot feathers have become textbook examples of sexual selection. Structural colors are crucial to avian biology, supporting functions such as communication and camouflage (*10*, *20*, *21*), and serve as models for bioinspired technologies (*22–24*).

Iridescence in bird feathers is a form of structural coloration that produces brilliant colors which vary with the angle of observation or illumination (*25*). Interest in these optical properties of avian plumage traces back to early studies by Robert Hooke and Isaac Newton (*26*, *27*). Iridescence originates within feather barbules, where multilayered nanostructures with high structural order are formed by materials of contrasting refractive indices (*10*, *25*, *28*). In birds, these nanostructures are formed from melanin-filled organelles called melanosomes (high refractive index) that are embedded in a matrix of keratin and air (lower refractive indices). This photonic lattice forms during feather growth and produces iridescent colors by reflecting light through constructive interference (*4*, *10*). Variation in melanosome morphology and arrangement, as well as barbule geometry, accounts for much of the nanostructural and optical diversity in iridescent plumage (*15*, *25*, *29–32*). Despite decades of optical and anatomical work clarifying the physical underpinnings of avian iridescence, the genetic and molecular mechanisms that generate the underlying nanostructures remain essentially unknown (*2*, *11*, *33*).

Here, we combined genetic, cellular, and optical approaches to uncover molecular mechanisms of iridescent color formation and diversification in birds. We leveraged the striking color variation that spontaneously arose during domestication of the Indian peafowl (*Pavo cristatus*), in which Mendelian mutations drive the gain, loss, and diversification of iridescence across the body, including in their iconic eyespots (**Fig. 1**). These color varieties, each a single mutational step from wild-type, offer a powerful system to link genotype, nanostructure, and optical phenotype. To determine whether similar mechanisms operate in natural populations, we profiled developing feathers from wild birds using single-cell transcriptomics.

Results

Peafowl color variation maps to melanogenesis genes

Our current understanding of the anatomical organization of photonic structures in avian feathers provides a framework to hypothesize mechanisms through which structural coloration might evolve in birds. Modifications in melanin biosynthesis, keratin production, feather morphogenesis, or spatial signaling pathways that coordinate intracellular organization are all potential mechanisms that might affect the resultant plumage color (*11*, *34*). However, progress in testing these hypotheses has been limited because iridescent traits rarely segregate within species, restricting the discovery of functionally relevant genetic variants (*35*, *36*). Domesticated color mutants of the Indian peafowl offer a rare opportunity to directly interrogate the molecular determinants of structural coloration.

To investigate the molecular basis of color variation in peafowl, we began by generating a *de novo* reference genome (**table S1; figs. S1 and S2**). Using long-read Nanopore sequencing, 10x Genomics linked-reads, and Hi-C chromatin conformation capture data, we assembled a chromosome-scale genome (1.04 Gb, N50 = 93.07 Mb) with 96.2% completeness of single-copy avian orthologs. Gene annotation, assisted by transcriptome data from five tissues, identified 21,618 protein-coding genes.

We then resequenced the genomes of 368 birds representing wild-type, color mutants known to be governed by single Mendelian loci, and color varieties likely arising from combinations of two or more mutations (mean depth = 10.1× ± 5.3 SD). We also profiled gene expression in regenerating feather follicles using RNA-sequencing for a subset of individuals (n = 27). Genome-wide association studies (GWAS) recovered significant loci for eight Mendelian mutations (**Fig. 2A**). Three had lead variants on the Z chromosome (*cameo*, *purple*, and *violet*), consistent with inheritance patterns observed by breeders. Each top candidate interval contained genes involved in distinct steps of the melanogenic pathway, including melanoblast differentiation (*EDNRB2*), regulation of melanogenesis (*MITF*), biogenesis and trafficking of melanosomes (*AP3S1*, *HPS3*, and *LYST*), melanosome transport to keratinocytes (*MLPH*), and melanin biosynthesis (*TYRP1*) (**Fig. 2B**).

Within each candidate region, we found putative causal variants, such as missense substitutions, premature stop codons, and exon deletions (**table S2**). In several cases, their functional relevance was supported by expression data or their overlap with evolutionarily conserved amino acid positions across avian species (**fig. S3**; **table S2**). Genotyping across our full cohort of samples revealed that four additional varieties result from combinations of these mutations (**table S2**), illustrating how epistatic interactions further expand color diversity in peafowl. Collectively, these results demonstrate that variation within melanogenesis genes alone can produce a wide range of effects on iridescent color phenotypes in bird feathers.

Photonic lattice geometry is highly sensitive to changes in melanogenesis

Next, we hypothesized that genetic variation in peafowl influences feather coloration by altering the organization of photonic nanostructures. To test this, we sampled six chromatic zones of the peacock eyespot (A1-A6; **Fig. 3A**) (*29*, *31*, *37*), and examined cross-sectional and longitudinal preparations of barbules from wild-type and the genetically characterized varieties using transmission electron microscopy (TEM). We quantified seven structural parameters that collectively describe the nanostructural organization of the barbule photonic lattice (**Fig. 3A; table S3**). We related ultrastructural features to optics via angle-dependent reflectance spectroscopy.

TEM imaging of wild-type feathers revealed that rod-shaped melanosomes formed a highly ordered photonic lattice, with variation in layer spacing and number generating the distinct hues across the eyespot (**Fig. 3A; fig. S4**), as previously documented (*29*, *30*, *37*). In contrast, color mutants exhibited pronounced deviations from the wild-type ultrastructure (**Fig. 3A-C; figs. S4 and S5**), and we applied linear discriminant analysis (LDA) to determine which features most strongly contribute to the differences (**Fig. 3D; tables S4 and S5)**. This analysis revealed clear separation between two main clusters along the first discriminant axis (LD1), which accounted for 90% of total variance (**Fig. 3D**). Melanosome length was the strongest contributor (Lm, $P = 6.1\times10^{-37}$) (**table S4**), which was inversely correlated with melanosome diameter (Dm) ($r = -0.76$), together defining the melanosome aspect ratio (rod-like versus spherical). *Bronze*, *cameo*, and *charcoal* mutants clustered on the negative LD1 axis, and TEM images revealed a collapse of the photonic lattice, with spherical and poorly ordered or disordered melanosomes replacing the aligned rods typical of wild-type iridescent feathers (**Fig. 3B and C; figs. S4 and S5**). Their reflectance spectra were flat or displayed weak features (**Fig. 3C; fig. S6**), indicating a loss of constructive interference. Our results thus reveal a gradient among peafowl mutants in which progressive melanosome rounding or irregularity corresponds to increasing lattice disorder (strongly iridescent types < *bronze* < *cameo* < *charcoal*) (**Fig. 3B and C**; **table S5**).

The *opal* variety exhibited a partially ordered lattice with fewer layers, yielding weaker iridescence (**Fig. 3B and C; figs. S4-S6**). *Violet*, *purple*, and *ultramarine* mutants retained well-organized multilayered lattices (**Fig. 3C**; **fig. S4**) and clustered close to wild-type (**Fig. 3D**). Within this iridescent group, LD1 (62.8% of variance) was driven by melanosome diameter (*Dm*; $r = 0.94$, $P = 1.2 \times 10^{-14}$), with additional contributions from inter-melanosome spacing (*b*; $r = 0.54$, $P = 0.002$) and inter-layer distance (*a*; $r = 0.38$, $P = 0.04$) (**Fig. 3D**; **table S4**). LD2 (25.2% of variance) was shaped mainly by layer number (*N*; $r = 0.81$, $P = 7.3 \times 10^{-8}$). Comparative analyses relative to wild-type highlighted the phenotype-specific shifts (**table S5**). *Opal* differed most strongly in layer number (*N*; $P = 1.1 \times 10^{-4}$), whereas *purple* and *violet* diverged mainly in melanosome diameter (*Dm*; $P = 2.8 \times 10^{-6}$) and in inter-melanosome spacing within layers (*b*; $P < 3.0 \times 10^{-2}$). *Ultramarine* showed differences only in inter-melanosomal spacing (*b*; $P = 2.0 \times 10^{-2}$). The number of layers (*N*) also differed significantly in *violet* ($P = 4.4 \times 10^{-2}$) and, although not significant overall, exhibited pronounced regional variation in *purple* and *ultramarine* (**figs. S4 and S5**).

These ultrastructural changes correlate directly with optical variation. A blue shift in peak reflectance from *purple* to *violet* to *ultramarine* (**fig. S6**) is associated with decreasing inter-melanosome spacing within layers, which shortens the optical path and shifts constructive interference toward shorter wavelengths. Across chromatic zones and iridescent varieties, layer number (*N*) was positively associated with color saturation (84%; $t = 5.7$, $df = 94$; $P = 1.3 \times 10^{-7}$), consistent with enhanced constructive interference in barbules with more layers (**fig. S7**). Thus, repetition (layer number) and periodicity (melanosome spacing) are tunable features of photonic lattices that are modulated by melanogenesis genes to control hue, saturation, and spectral bandwidth in iridescent plumage. Our findings also reveal a continuum of nanostructural integrity among peafowl color variants and suggest a hierarchical program underlying iridescent color production: some mutations act as disruptive switches that collapse lattice formation, and others function as quantitative modulators that fine-tune color by adjusting lattice geometry.

Cellular and biochemical basis of iridescence

To elucidate how cellular and biochemical mechanisms shape these nanostructural patterns in peafowl feathers, we analyzed regenerating feather follicles and mature neck feathers from wild-type and mutant birds. Longitudinal follicle sections were examined by light microscopy to assess melanin distribution in melanocyte and keratinocyte layers, as well as the melanocyte dendritic branching (**Fig. 4A**), which mediates melanosome transfer into keratinocytes that become the feather's structural matrix. In parallel, we performed chemical analysis of mature feathers to evaluate how mutations affect melanin biosynthesis and deposition (**Fig. 4B**; **table S6**). We focused on neck feathers because they regenerate independently of the molting cycle and exhibit uniform coloration, facilitating comparisons of cellular and biochemical phenotypes.

In wild-type follicles, melanosomes were abundant and efficiently transferred from melanocytes to keratinocytes through dendritic projections, producing a uniform pigment distribution (**Fig. 4A**). Biochemically, wild-type feathers had a high eumelanin-to-pheomelanin ratio (~10:1; **Fig. 4B**). The weakly iridescent (*bronze*) and non-iridescent mutants (*cameo* and *charcoal*) showed diverse cellular and biochemical changes. In *charcoal* follicles, dendrites and melanosome transfer were present, but melanosomes formed aggregates within melanocytes and keratinocytes (**Fig. 4A**). Despite appearing visually darkest, *charcoal* contained 67% less total melanin than wild-type feathers ($t = 8.4$, $df = 6$, $P = 1.6 \times 10^{-4}$) and comparable levels to *violet* and *purple* (**Fig. 4B**), which appear far more vibrant, illustrating a disconnect between pigment abundance and perceived brightness arising from structural organization. *Cameo* and *bronze* displayed disrupted melanin distribution, with clustered pigment deposits within keratinocytes indicating melanosome aggregation (**Fig. 4A**). In *cameo*, pigment-free regions within melanocytes suggested the presence of large intracellular vacuoles. Importantly, across all three mutants, eumelanin content was markedly reduced, but the pheomelanin levels remained almost unaltered (**Fig. 4B**), hence reducing the eumelanin-to-pheomelanin ratio (~1.5:1–4:1). These compositional shifts coincided with the transformation of rod-shaped melanosomes into spherical forms (**Fig. 3B and C**). Although the underlying genes act at distinct steps of melanogenesis, none are known to directly regulate melanosome morphology. The repeated emergence of spherical melanosomes instead suggests that altered melanin composition influences melanosome shape—either directly or by perturbing normal melanosome maturation—which ultimately compromises the integrity of the photonic lattice.

The *opal* mutant showed a distinct cellular phenotype, characterized by melanocytes densely filled with melanin but lacking visible dendritic projections to keratinocytes, which receive few melanosomes (**Fig.**

4A). This suggests disrupted melanosome transfer. Biochemically, *opal* feathers exhibited a large reduction in total melanin (–73%; $t = 11.4$, $df = 7$, $P = 8.9 \times 10^{-6}$; **Fig. 4B**), but unlike *cameo*, *bronze*, and *charcoal*, this decrease affects both eumelanin and pheomelanin (**Fig. 4B**). This more balanced loss appears to preserve the rod-shaped melanosomes typical of wild-type, in contrast to the spherical forms that characterize other low-melanin mutants. Consequently, despite melanin levels comparable to *charcoal*, the retention of rod-shaped melanosomes permits partial nanostructural organization, making *opal* appear much lighter and vivid.

Hue-shifted iridescent variants (*purple*, *ultramarine*, and *violet*) showed feather follicles with cellular organization comparable to wild-type, with only slight differences (**Fig. 4A**). *Violet* and *purple* follicles contained a higher proportion of light brown melanosomes, indicative of delayed or incomplete maturation as expected from *TYRP1* mutations (*38*), whereas *ultramarine* follicles exhibited dense pigmentation. All three phenotypes maintained wild-type-like eumelanin-to-pheomelanin ratios (*purple*: 9:1; *ultramarine*: 11.4:1; *violet*: 8.5:1) and retained the characteristic rod-shaped melanosomes (**Fig. 4B**).

Despite these similarities, these mutants fall into two distinct mechanistic categories. *Purple* and *violet* feathers showed large reductions in total melanin content (*purple*: –50%, $t = 6.7$, $df = 6$, $P = 5.3 \times 10^{-4}$; *violet*: –72%, $t = 9.4$, $df = 6$, $P = 8.5 \times 10^{-5}$) (**Fig. 4B**), yet sufficient to support ordered multilayered structures (**Fig. 3C**). Both mutants also exhibited reduced inter-melanosome spacing (see above). Because *TYRP1* encodes a core enzyme in eumelanin synthesis and contributes to melanosome maturation rather than spatial organization (*38*), it is unlikely to directly regulate inter-melanosome packing. Instead, the packing changes are most plausibly an indirect consequence of the increased melanosome aspect ratios resulting from reduced melanosome diameter (*Dm*) in both mutants compared to wild-type (**Fig. 3B**). This is consistent with known effects of *TYRP1* on melanosome morphology (*39*). Thinner melanosomes likely promote more efficient packing because the increased melanosome aspect ratio facilitates tighter alignment along their long axes. Thus, variation in melanosome aspect ratio represents an additional feature of melanogenesis that can influence photonic assembly and optical performance. The observed differences between *purple* and *violet* feathers likely reflect an additional layer of interaction between melanosome abundance and melanosome shape.

In contrast, we found that *ultramarine* feathers, likely caused by an *MITF* mutation, showed no significant change in pigment abundance (+0.06%, $t = 4.9 \times 10^{-3}$, $df = 5$, $P = 0.1$) (**Fig. 4B**), despite intense cellular pigmentation in the follicle sections (**Fig. 4A**). Analyses of melanosome morphology and pigment chemistry likewise revealed no detectable differences from wild-type (**Fig. 4B**; **fig. S5**), indicating that pigment quantity or composition cannot account for the observed optical properties. Our ultrastructural data show that the *ultramarine* photonic lattice is densely packed, with significantly reduced inter-melanosome distance (*b*) (**Fig. 3C; figs. S4 and S5**). TEM images further indicated a unique organizational anomaly in *ultramarine*, which is absent in wild-type and other mutants, consisting of clusters of disorganized melanosomes near the basal boundary of the photonic crystal and occasionally dispersed melanosomes within the medullary region of the barbule (**fig. S8**). The anomaly suggests a developmental mechanism in which photonic lattice assembly was prematurely arrested and melanosomes failed to integrate into the lattice within the appropriate temporal window. A loss of temporal coordination between melanosome delivery and lattice formation could also underlie the observed alterations in periodicity, possibly reflecting melanosome deposition during phases of altered keratin matrix viscosity. This interpretation aligns with several known functions of *MITF*, including its roles in regulating melanocyte differentiation and the timing of melanosome maturation (*40*). Taken together, our biochemical, ultrastructural, and photonic analyses show that nanostructure formation is a plastic trait responsive to melanosome availability and morphology as well as to the temporal dynamics of pigment deposition during feather development.

Single mutations in melanogenesis genes can induce iridescent plumage

The *black-shoulder* mutation confers a localized gain of iridescence, unlike the other Mendelian traits that alter existing iridescent effects. It replaces the light-dark barred pattern on the shoulders and upper back of wild-type peacocks, regions mostly lacking iridescence, with vivid green and blue hues. Other regions of the plumage appear unaffected (**Fig. 5A**), providing an opportunity to investigate the molecular and developmental basis of newly evolved iridescent traits. To identify the causal variant, we conducted a GWAS comparing *black-shoulder* ($n = 93$) and wild-type birds ($n = 256$, scored for wing color phenotype). This analysis revealed a significant association ($P = 5.0 \times 10^{-28}$) overlapping the *MC1R* gene (**Fig. 5B; fig. S9**), which encodes a receptor that regulates melanin synthesis via the cAMP pathway (*41*).

We next assessed gene expression in regenerating wing covert feather follicles from *black-shoulder* ($n = 4$) and wild-type ($n = 6$) birds using RNA-seq. *MC1R* expression was similar between phenotypes (**Fig. 5C**), suggesting that functional differences arise from coding changes rather than changes in gene regulation, despite the patch-specific effects. Examination of the *MC1R* coding sequence revealed two mutations: a missense mutation (the GWAS lead variant) and a frameshift mutation (**table S2**). Genotyping showed that *black-shoulder* individuals were either homozygous for one mutation or compound heterozygous, indicating genetic heterogeneity underlying the same phenotype. The frameshift likely produces a loss-of-function allele, and the similar phenotype caused by the missense variant suggests both impair *MC1R* function. Allelic imbalance tests in heterozygote peacocks showed no expression bias (**Fig. 5C**), reinforcing that the phenotype arises from altered protein function.

To understand the biochemical and photonic basis of this gain of iridescence, we analyzed melanin content and feather ultrastructure. Chemical analyses showed a ~1.7-fold increase in eumelanin in *black-shoulder* wing covert feathers relative to the wild-type ($t = 12.3$, $df = 3$, $P = 4.0 \times 10^{-5}$), whereas eumelanin in neck feathers and pheomelanin in either region showed no significant difference ($t < 0.4$, $df = 3\text{-}5$, $P > 0.4$; **Fig. 5D**). These results, together with previous work in birds (*42*, *43*), suggest that *MC1R* loss-of-function mutations have phenotypic effects distinct from those in mammals, where loss of *MC1R* typically increases pheomelanin production (*44*). TEM imaging showed organized rod-shaped melanosomes in dark bands of wild-type feathers, but light bands were sparsely pigmented and lacked discernible structure (**Fig. 5E**). In *black-shoulder* feathers, however, melanosome density and layering were often increased toward feather tips, and the multilayered organization extended into regions corresponding to the wild-type light bands. Both features account for the enhanced photonic properties of *black-shoulder* feathers (**Fig. 5E**). These findings suggest that the *MC1R* mutations promote melanin production or deposition to a threshold that allows the self-assembly of rod-shaped melanosomes into complex multilayers, even in feather barbs previously devoid of such structures. Thus, peafowl and possibly other birds possess a latent capacity to form ordered photonic nanostructures that can be activated by single mutations enhancing melanogenesis.

Transcriptomic drivers of iridescence in nature

To investigate iridescent color formation in wild birds, we performed single-cell RNA sequencing (scRNA-seq) on regenerating feathers of the emerald dove (*Chalcophaps indica*) and mallard duck (*Anas platyrhynchos*). These species represent two additional avian orders and develop asymmetric feathers in their wings, where iridescent and non-iridescent barbules differentiate on opposite sides of the same feather (**Fig. 6A**). TEM confirmed ordered nanostructures within iridescent barbules and their absence in non-iridescent ones (**Fig. 6A**; **fig. S10**). To identify molecular drivers of these differences, we microdissected regenerating follicles to isolate the sides expressing iridescent and non-iridescent colors and generated eight scRNA-seq libraries (two replicates per barbule type per species). This within-feather design enabled direct comparisons of cell composition and gene regulation between barbule types.

We identified 31,204 high-quality cells representing all major cell types in growing feathers (**Fig. 6B; figs. S11-14**). Melanocytes formed two transcriptionally distinct clusters corresponding to eumelanin- (melanocytes 1) and pheomelanin-producing (melanocytes 2) subtypes (see Methods). Using these cell populations, we contrasted iridescent and non-iridescent barbules to identify both differentially expressed genes (DEGs) and differential cellular composition. In the emerald dove, melanocytes exhibited the most noteworthy changes in both cellular composition and transcriptional profile (**Fig. 6C**). They were threefold more abundant in iridescent barbules (26.7% vs. 8.6%; $P = 1.7 \times 10^{-2}$) and showed the largest number of DEGs. This expression divergence was independent of cell-type abundance and more pronounced within the eumelanin-producing subtype. In mallard, we saw no consistent compositional differences, yet melanocytes again showed the largest transcriptional divergence (**Fig. 6C**). We validated these findings with histological analysis of regenerating feathers and pigment quantification in mature feathers, which confirmed increased cellular pigmentation and greater melanin deposition in the iridescent side (**Fig. 6A and D**). These results suggest that cellular and transcriptional changes in melanocytes are central to iridescent color development in wild birds, consistent with our findings in peafowl.

Because melanocytes showed the strongest transcriptional divergence, we first focused on their DEGs (**Data S1 and S2**). When the melanocyte subclusters were analyzed either separately or together, only eight genes were shared between both species (*CCZ1*, *CHN1, DCLK1*, *DCX*, *DLC1*, *FTH1*, *GPNMB*, and *NCALD*), indicating that iridescent coloration is produced by distinct gene sets in each lineage despite their convergent

nanostructural architecture and optical function. None of the DEGs encoded melanin-synthesis enzymes; instead, expression changes implicated melanocyte signaling, ion homeostasis, melanosome dynamics, and dendritic transfer. In the emerald dove, v-ATPase subunits (*ATP6V1E1*, *ATP6V0A1*), the cation channel *TPCN2*, and the potassium channel *KCNJ13* showed differential expression, consistent with active tuning of melanosome pH and ion balance (processes essential for melanogenesis) in iridescent barbules. In mallard, *cAMP*-responsive effectors downstream of *MC1R* (a DEG in the dove) were elevated (*RAPGEF4* and *ARPP21*), along with trafficking and autophagy regulators (*UVRAG*) and cytoskeletal/adhesion genes involved in dendrite elaboration and vesicle transport (*TMSB4X*, *CHN1*, *KANK1*, *LPP*). Together with increased *GPNMB* and *CCZ1* expression (DEGs shared between both species), these signatures point to more active melanosome maturation and intracellular trafficking as a common feature of iridescent barbule formation. Intercellular signaling also shifted in iridescent feathers and was reflected in gene ontology (GO) enrichment (**fig. S15**): genes associated with GTPase–mediated and ephrin-receptor signaling were elevated in mallard iridescent barbules (*RAPGEF4, CHN1, EPHA7*); WNT-axis components were reduced in both species (*DAAM2* in mallard; *LGR4* in dove); and integrin signaling decreased in dove (*ITGA4*). Three of the genes shared by both species (*CHN1*, *DCLK1*, *DCX*) and several top-10 DEGs (*NBEA*, *VAT1*/*VAT1L*) were linked to neuronal and synaptic processes (GO analyses also confirm an enrichment of many neuron/synapse-related terms; **fig. S15**), raising the hypothesis that synaptic-like vesicular processes contribute to differences in melanosome transfer along melanocyte dendrites between barbule types. These data reveal extensive rewiring of melanocyte biology between iridescent and non-iridescent barbules, while expression of the core enzymatic machinery of melanogenesis remains unchanged.

To assess gene regulation beyond melanocytes, we examined DEGs in keratinocytes and mesenchymal cells (**Data S1 and S2**). Many were functionally linked to melanogenesis through their regulatory effects on melanocyte behavior (*45*, *46*), indicating that altered cross-talk between pigment cells and neighboring tissues distinguishes iridescent from non-iridescent barbules. Among the 15 keratinocyte DEGs with well-characterized proteins, six have known connections to melanocyte regulation, including paracrine factors upregulated in iridescent barbules that enhance melanocyte recruitment and proliferation (*KITL* and *NRG1* in mallard) and adhesive molecules mediating keratinocyte-melanocyte communication (*COL17A1* in dove). Mesenchymal cell DEGs likewise showed transcriptional signatures of differential paracrine regulation with secreted factors known to be implicated in melanocyte differentiation, migration, and pigment output, including *NDP* and *BMP2* in dove and *ASIP* and *FST* in mallard. Finally, several DEGs may contribute to iridescence through effects on the keratin matrix and tissue mechanics (*11*). For example, keratinocyte DEGs included feather structural components encoded by keratin genes that were consistently downregulated in iridescent barbules (*KRT1L*, *KRTβL*, and *KRT6A*). Taken together, our findings point to a hierarchical regulatory network centered on melanogenesis, but distributed across tissues, that coordinates the assembly of feather photonic structures in nature (**Fig. 6E**).

Discussion

The formation of iridescent coloration requires an astonishing degree of structural precision. By analyzing peafowl color variants, we connect genetic and biochemical mechanisms to specific macroscopic optical properties, showing how mutations in melanogenesis genes alter the composition, shape, and self-organization of photonic nanostructures, producing color shifts from subtle hue changes to complete loss of iridescence. Single-cell analyses from wild species further underscore melanogenesis as a major contributor to structural color and reveal a broad network of transcriptomic changes associated with iridescence. Although other cellular processes can theoretically influence feather nanostructures (*11*, *47*, *48*)—and our single-cell data reveal some signatures of these processes (e.g., changes in keratin gene expression)—melanogenesis appears as a central driver of iridescent color evolution. We propose that melanogenesis is especially effective at tuning iridescence because its multistage pathway, governed by dozens to hundreds of interacting factors (*44*), forms a broad mutational target in which even minor perturbations can induce major nanostructural and optical shifts. Moreover, because mutations in melanogenesis genes often have low pleiotropic effects, and because these genes are frequently active in peripheral tissues such as the skin, functional changes can evolve with fewer deleterious consequences (*44*). In contrast, modifications to keratinization or feather morphology are likely more constrained by their mechanical and insulative functions. These properties of melanogenesis may help explain why iridescent colors are so widespread and have evolved repeatedly across avian lineages (*19*).

In contrast to comparative studies of distantly related species, our analysis leverages the shallow genetic

divergence among peafowl varieties that exhibit cellular, nanostructural, and optical changes arising from single mutational events. The genes identified span multiple stages of melanogenesis. Despite this functional diversity, our histological and biochemical data show that these genes control the manifestation of iridescence by altering melanin composition, particularly the balance of pheomelanin and eumelanin. In all cases in which eumelanin declined and pheomelanin remained largely unchanged, rod-shaped melanosomes transitioned into spherical forms. This association suggests that changes in pigment composition can influence melanosome morphology without additional genetic input directly controlling organelle shape or architecture (*49*, *50*). In turn, we show that the presence of spheroid melanosomes disrupts photonic lattice organization, collapsing multilayered nanostructures into haphazard arrangements that produce black or brown plumage. These observations provide molecular support for models in which photonic lattices emerge primarily through self-assembly processes, such as depletion attraction and other short-range interactions, decoupled from direct cellular control (*47*). This mechanistic link also explains why, across diverse bird species, pheomelanin-rich feathers rarely display iridescence (*11*, *25*) and how shifts in melanin composition can rapidly drive transitions between iridescent and non-iridescent states.

Once photonic lattices form, melanin abundance and subtle variations in rod-like melanosome morphology fine-tune reflected wavelengths and color intensity by altering lattice periodicity and stacking. Our observations in the *ultramarine* variant further suggest that the temporal coordination of melanosome delivery and lattice assembly may critically influence photonic effects. These findings imply that key elements of feather nanoarchitecture respond dynamically to the cellular and biochemical environment, rather than being rigidly encoded by genetic positional blueprints that specify layer number or spacing. Nevertheless, developing cells can actively regulate melanin production, thereby altering the optical outcomes of self-assembly and driving diversification of hues and other aspects of iridescent coloration. In this context, the *black-shoulder* mutation demonstrates how a single change that elevates melanin deposition can convert non-iridescent into iridescent plumage by activating multilayered photonic structures in regions that previously exhibited minimal optical effects, suggesting that this remarkable coloration mechanism may be more evolutionarily accessible than previously thought. Although the evolution of iridescence is often viewed as a gradual increase in optical complexity (*25*, *29*, *34*), our findings show that species already capable of forming multilayered nanostructures have a latent capacity to evolve novel iridescent patterns without intermediate steps through relatively minor adjustments in melanin synthesis.

In summary, this study provides molecular insights into the generation of iridescent colors of bird feathers, a phenomenon that was previously understood mainly from physical and anatomical perspectives. More broadly, it illustrates how complex biological structures and functions emerge from the interplay between genetically encoded developmental programs, which establish or constrain the conditions for self-organization, and universal self-organizing forces that give form and function to those programs.

METHOD DETAILS

Ethics statement

All experimental procedures involving animals were carried out in compliance with Directive 2010/63/EU on the protection of animals used for scientific purposes. These procedures were reviewed and approved by the Animal Welfare and Ethics Committee of CIBIO/BIOPOLIS (approval reference: 2018_01 and 2019_03).

Genome sequencing of the Indian peafowl

Sampling and DNA isolation. A draft reference genome sequence of the Indian peafowl (*Pavo cristatus*) was generated from a phenotypically wild-type male. Whole blood was collected using a sterile needle and transferred into a heparin-free capillary. The sample was snap-frozen in liquid nitrogen and stored at -80°C until DNA extraction. High-molecular-weight genomic DNA was isolated from 5 µl of blood using a modified salt-based protocol (*51*), followed by a purification step using AMPure XP magnetic beads (Beckman Coulter) at a 3× beads-to-sample ratio. DNA quantity, purity, and integrity were assessed with a NanoDrop spectrophotometer, a Qubit dsDNA BR Assay Kit (ThermoFisher Scientific), and Agilent Genomic DNA ScreenTape (Agilent).

Nanopore long-read sequencing. Sequencing libraries were prepared using the Oxford Nanopore Technologies (ONT) 1D Sequencing Kit (SQK-LSK109). Briefly, 1.0 µg of high-molecular-weight genomic DNA underwent DNA repair and end-repair using the NEBNext FFPE DNA Repair Mix (NEB) and the NEBNext UltraII End Repair/dA-Tailing Module (NEB). Sequencing adaptors were ligated, and the libraries were purified using 0.4× AMPure XP Beads (Beckman Coulter) and eluted in the provided buffer (SQK-LSK109). Sequencing was performed on a GridION Mk1 (ONT) using R9.4.1 FLO-MIN106D flow cells, with data collected over 72 hours. Run quality parameters were monitored in real time using MinKNOW (v4.2.5).

Data from eight flow cells were base-called using *Guppy* (v3.2.6+afc8e14 for seven flow cells; v3.2.10+aabd4ec for one flow cell). Reads were filtered to remove those with: (1) an average Phred quality score ≤7; (2) lengths <1 kb; (3) low complexity sequences exceeding 40%; or (4) matches to the control sequence (lambda phage, 3.5 kb). Additionally, contamination was assessed using *Kraken* (v0.10.5-beta) (*52*), and reads matching *Spirometra erinaceieuropaei*, Human gammaherpesvirus-4, cloning vector pCA-DEST2303, or *Mycoplasma hyorhinis* (the four contributing to ~0.01% of reads) were excluded. After filtering, 4,599,390 ONT reads totaling 43.2 Gb (33× coverage) with an N50 of 16.6 kb were retained for downstream analysis.

Linked-read sequencing. The linked-read library was prepared using a Chromium Controller (10x Genomics) and the Genome Reagent Kits v2 (10x Genomics), following the manufacturer's protocol. In brief, 10 ng of high-molecular-weight genomic DNA was partitioned in GEM (Gel Bead-in-Emulsion) reactions containing unique barcodes (Gemcode) and loaded onto a Chromium controller chip. The droplets were recovered, isothermally incubated, and fractured. The intermediate DNA library was then purified and size-selected using Silane and Solid Phase reverse immobilization (SPRI) beads. An Illumina-compatible paired-end sequencing library was prepared following 10x Genomics recommendations and validated using an Agilent 2100 BioAnalyzer with the DNA 7500 assay (Agilent). Sequencing was performed on an Illumina NovaSeq 6000 platform using paired-end reads (2 × 151 bp).

The resulting reads were preprocessed for scaffolding using *LongRanger Basic* (v2.2.2), which performed basic read and barcode processing, including read trimming (removing the first 16 base pairs of read one if they match a valid barcode), barcode error correction, barcode whitelisting, and attaching barcodes to read IDs (producing barcoded reads). Additionally, the original linked read data was processed for use as a standard paired-end library during assembly. All reads were filtered for contaminants by mapping with *gem-mapper* (*53*) (allowing up to 2% mismatches) against a contamination database that included phiX, UniVec sequences, *E. coli*, the

complete mitochondrial sequence of Indian peafowl (NC_024533.1), and the four contaminants detected with *Kraken* (as described above). A total of 262,153,402 paired reads, corresponding to 78.6 Gb (~60× coverage), successfully passed the quality filters and were used for genome assembly.

Chromatin conformation capture sequencing (Hi-C). A Hi-C library was prepared using the Arima High Coverage Hi-C kit (Arima Genomics), following the manufacturer's instructions. Sequencing was performed on an Illumina HiSeq 4000 platform using paired-end reads (2 × 151 bp). A total of 317,826,943 pairs of reads were aligned against the draft assembly using *Juicer* (v1.6) (*54*). This process yielded 261,213,953 Hi-C contacts, which were used for subsequent scaffolding.

Contig construction and scaffolding. The Indian peafowl reference genome sequence was assembled using a combination of data sources and multiple iterative steps.

1. Hybrid assembly:
   A hybrid genome assembly was generated with *MaSuRCA* (v3.4.1) (*55*, *56*), using the Illumina linked-read data (60× coverage, treated as paired-end reads) and the pre-processed Nanopore reads (33× coverage) to create mega-reads. These mega-reads were assembled with *Flye* (v2.5) (*57*). Default parameters were used in *MaSuRCA*, with adjustments to the following settings: USE_LINKING_MATES = 0, MEGA_READS_PASS=0, JF_SIZE = 26200000000, and FLYE_ASSEMBLY=1.
2. Correction and scaffolding with linked reads:
   Barcoded 10x linked reads were used to correct and scaffold the assembly. Misassemblies were identified and resolved with *Tigmint* (v1.1.2) (*58*), followed by scaffolding with *ARKS* (v1.0.3) (*59*) and *LINKS* (v1.8.5) (*60*), as per a published pipeline (http://protocols.faircloth-lab.org/en/latest/protocols-computer/assembly/assembly-scaffolding-with-arks-and-links.html).
3. Hi-C scaffolding:
   The genome was further scaffolded using the Hi-C data with *Juicer* (v1.6) and the *3D de novo assembly (3D-DNA)* pipeline (v180114) (*61*).

Curation of the chromosome-level assembly. The *JuiceBox Assembly Tools* (*JBAT*) (v1.11.08) (*62*) were used for visual inspection and revision of the contact map. The putative misassemblies were revised and corrected. The decisions were made by taking into account breaks in synteny against the chicken chromosomes (GRCg6a) based on a pairwise alignment obtained with *nucmer* (*63*) and inspected with *Dot* (https://github.com/dnanexus/dot). Small scaffolds potentially containing inversions or translocations were examined within the super-scaffolds. Inversions or translocations detected in the dot plot were manually corrected using JBAT, ensuring changes maintained or improved consistency with the Hi-C contact map. Despite the ~35 million years of divergence between chickens and Indian peafowl, broad synteny across the majority of the genome is expected. However, intra-scaffold rearrangements consistent with the Hi-C contact map were left unchanged, as they could represent real species-specific differences. After completing the review, the assembly was finalized by exporting the modified assembly files and generating new FASTA files using the *run-asm-pipeline-post-review.sh* script from the *3D-DNA* pipeline. Finally, this revised assembly was sealed and closed by placing the debris scaffolds into gaps by running *run-asm-pipeline-post-review.sh* with options -s seal and -i 15000. The completeness of the final genome assembly was assessed using *BUSCO* (v5.8.2) (*64*) and the 8,338 single-copy avian ortholog genes in the *aves_odb10* database. Summary statistics for the final assembly are provided in **table S1** and **figs. S1 and S2**.

Genome annotation of the Indian peafowl

RNA sequencing. RNA-seq data from five tissues of the same male individual used for the reference genome sequence were generated to support gene annotation. Total RNA was extracted from brain, lung, testes, skin, and kidney tissues and quantified using the Qubit RNA BR Assay Kit (Thermo Fisher Scientific). RNA integrity was assessed with the Agilent RNA 6000 Pico Kit on a Bioanalyzer 2100 (Agilent). RNA-seq libraries were prepared using the KAPA Stranded mRNA-Seq Kit for Illumina platforms (Roche), following the manufacturer's protocol with 500 ng of total RNA as input. Library size and quality were evaluated using the High Sensitivity DNA Bioanalyzer Assay (Agilent). The RNA-seq libraries were run on an Illumina HiSeq 4000 platform using paired-end reads (2 × 101 bp).

Repeat identification. Repeats in the reference genome assembly were annotated using *RepeatMasker* (v4.0.7) (http://www.repeatmasker.org) with a custom repeat library available for chicken.

Gene model inference. Gene annotation was performed by integrating transcript alignments, protein alignments, and *ab initio* gene predictions.

1. Transcript alignments based on RNA-seq data:
   The multi-tissue RNA-seq reads were aligned to the genome using *STAR* (v.2.7.2a) (*65*) and transcript models were built for each tissue separately with *Stringtie* (v2.0.1) (*66*). These models were combined using *TACO* (*67*), and curated junctions to be used during the annotation process were identified with *Portcullis* (*68*) after mapping with *STAR*. *PASA* (v2.3.3) (*69*) was used to assemble transcripts, incorporating 2,104 Indian peafowl nucleotide sequences retrieved from NCBI (July 2020). Coding regions in the transcripts were identified with *TransDecoder program* (https://github.com/TransDecoder/TransDecoder), part of the *PASA* package.
2. Protein alignments:
   Complete chicken, human, and turkey proteomes were downloaded from UniProt (July 2020) and aligned to the genome using *spaln* (v2.4.7) (*70*).
3. *Ab initio* predictions:
   Ab initio gene predictions were performed on the repeat masked assembly with three different methods: *GeneID* (v1.4), Augustus (v3.2.3) (*71*), and *Genemark-ES* (v2.3e) (*72*) with and without RNA-seq junction evidence. The gene predictors were run with trained parameters for the human genome except for *Genemark* that runs in a self-trained manner.
4. Consensus gene models:
   EvidenceModeler (EVM; v1.1.1) (*73*) integrated all transcript, protein, and *ab initio* data to produce consensus coding DNA sequence (CDS) models. Untranslated regions (UTRs) and alternative splicing forms were annotated through two rounds of *PASA* updates.
5. Functional annotation:
   Functional annotation of the proteins was conducted using *Blast2GO* (*74*). First, a *BLASTp* (*75*) search was performed against the *nr* database (accessed February 2021). Second, protein domains were identified with *InterProScan* (*76*). The results were combined in *Blast2GO* to generate the final functional annotations.

ncRNA annotation. Non-coding RNAs (ncRNAs) were annotated through the following steps. First, the *cmsearch* tool (v1.1; from *Infernal*) (*77*, *78*) was used to scan the genome against the RFAM RNA family database (v12.0) (*79*). Transfer RNAs (tRNAs) were identified with *tRNAscan-SE* (v1.23) (*80*). To detect lncRNAs, *PASA* assemblies not annotated as protein-coding genes were screened to identify expressed, non-translated sequences. Assemblies longer than 200 bp and not covered >80% by small ncRNAs were classified as lncRNAs. lncRNA transcripts

were clustered into genes based on shared splice sites or significant sequence overlap. Summary statistics of the genome annotation are given in **table S1**.

Whole-genome resequencing and genetic mapping

Sampling of peafowl color varieties and DNA isolation. Blood samples or growing feathers were collected from 352 captive-raised Indian peafowl for whole-genome sequencing (**table S7**). Each individual was scored for the body color and wing phenotype, since they segregate independently. Color phenotypes were scored based on visual inspection, sometimes supplemented by genotype information provided by breeders for individuals known to be heterozygous for specific recessive mutations based on crossing records and studbook data. Samples included wild-type individuals and 20 distinct color varieties. Nine of these varieties are known to result from Mendelian mutations and therefore are determined by a single locus in the genome: *bronze*, *cameo*, *charcoal*, *midnight*, *opal*, *purple*, *ultramarine*, *violet*, and *white*/*pied*. From segregation analyses in crosses, three of these varieties are known to be sex-linked and located on the Z-chromosome (*cameo*, *purple,* and *violet*). For the remaining varieties (*brown*, *chestnut*, *hazel*, *indigo*, *ivory*, *jade*, *peach*, *Prussian-blue*, *raw-umber*, *Sonja violet*, and *taupe*), the inheritance patterns are either unknown, poorly supported, or attributed to combinations of two or more mutations. To investigate and take into account potential introgression from green peafowl (*Pavo muticus*) that might interfere with genetic mapping, four samples from green peafowl and 12 *Spalding* individuals, hybrids between Indian and green peafowl, were also collected.

Whole blood was collected as described earlier, but stored in 96% ethanol until DNA extraction. For birds where blood collection was not feasible, naturally molting feathers were plucked and similarly stored in ethanol. Genomic DNA was extracted from blood samples using a modified salt-based protocol, and from feathers it was extracted using the QIAamp UCP DNA Micro Kit (Qiagen). Potential RNA contamination was removed by treating DNA extracts with RNase-A (Roche). Following DNA isolation, DNA quality and purity were assessed using spectrophotometry (Nanodrop) and fluorometric quantitation (Qubit dsDNA BR Assay Kit, ThermoFisher Scientific).

Illumina sequencing and read mapping. Genome-wide polymorphism data for genetic mapping was generated using whole-genome Illumina sequencing. Sequencing libraries were prepared for each of the 368 individuals (Indian peafowl, green peafowl, and their hybrids). For blood samples, which have higher DNA quality and larger quantities, libraries were constructed using the TruSeq DNA PCR-Free Library Preparation Kit (Illumina). For some feather follicle samples, which often yield lower DNA quality or quantity, libraries were prepared using the PCR-based Nextera XT Library Preparation Kit (Illumina). The quality of each library was assessed based on fragment size distribution, evaluated on an Agilent 2200 TapeStation using an HS D5000 Screen Tape (Agilent Technologies), and molarity, calculated using a KAPA qPCR library quantification kit (Roche). Libraries were sequenced using paired-end reads (2 × 150 bp) on an Illumina instrument.

Prior to data analyses, sequencing read quality was inspected with *FastQC* (v0.11.8) (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads were then aligned to the Indian peafowl genome assembly generated in this study using *BWA-MEM* (v0.7.17-r1188) (*81*) with default parameters. Duplicate reads were flagged for downstream analysis using the *MarkDuplicates* function in *Picard* (v3.0.0) (http://broadinstitute.github.io/picard). Read group information was added to each BAM file using the *Picard* function *AddOrReplaceReadGroups*. Sequencing and mapping summary statistics were calculated using *SAMtools* (v1.11) (*82*). After mapping and duplicate removal, the final effective mean sequencing depth was 10.1× ± 5.3 (**table S7**).

Variant discovery and genotype calling. Single nucleotide polymorphisms (SNPs) and small insertions or deletions (indels) were identified using *GATK* (v4.2.6.1) (*83*). Briefly, *gVCF* files were generated for each individual using the *HaplotypeCaller* function requiring a minimum mapping quality of 30 and setting the heterozygosity value to 0.004. These files were subsequently combined into a single file using the function *CombineGVCFs*, and variants were called using the function *genotypeGVCFs* using default parameters except for the heterozygosity value (--heterozygosity 0.004). Prior to downstream analyses, additional filtering steps were applied at both the genotype and variant levels using *VCFtools* (v0.1.16) (*84*). The genotype of each individual was coded as 'missing data' if its quality was below 30 (--minGQ 30), or if its coverage was below 4× or higher than 58× (i.e., twice the average coverage of the individual with the highest coverage; --minDP 4 --maxDP 58). Following genotype filtering steps, variants with 50% or more missing genotype data across individuals (--max-missing 0.5 in *VCFtools*) were excluded.

Variant effect prediction. Filtered variants were annotated for their potential impact on protein-coding regions, including nonsynonymous, frameshift, splicing, and stop-gain mutations, using the *SnpEff* toolbox (v4.3t) (*85*). Based on these annotations, a subset of variants classified as high or moderate impact mutations was extracted for independent analysis of coding region variation. This filtering process resulted in a dataset of 228,553 variants.

Genome-wide association analysis. Before conducting the genome-wide association analysis, missing genotypes were imputed, and genotype data were phased using the software *Beagle* (v5.1) (*86*, *87*). Association testing for each marker was performed with *GEMMA* (v0.98.1) (*88*) using a linear mixed model and coding each color morph as a binary trait (mutant vs. all others). The varieties that were known or discovered to be combinations of mutations were sometimes excluded from the control or case groups. Individuals belonging to the *white* variety (n = 17) were excluded from the mapping analyses of the other varieties because the mutation is epistatic, masking all other phenotypes. To minimize the potential confounding effects of population stratification and relatedness among individuals, a kinship matrix was included in the model as a random effect. The kinship matrix was estimated using *GEMMA*'s centered relatedness matrix option (-gk 1). Additionally, the first three principal components (PCs) from a principal components analysis (PCA), calculated with *PLINK* (v1.90b6.26) (*89*), were included as covariates. Variants with a minor allele frequency (MAF) lower than 10% were not considered. Significance thresholds for association testing were determined using a Bonferroni correction ($-\log_{10}$ [0.05/ number of loci]) applied to the likelihood ratio test for allele frequencies from *GEMMA*. Manhattan plots summarizing the associations were generated with the *qqman R* package (*90*).

Bulk RNA-sequencing

Sampling and RNA extraction. RNA-seq data were generated from 37 regenerating feather follicle samples of Indian peacocks, including neck feather follicles from wild-type and mutant varieties, and wing feather follicles from wild-type and *black-shoulder* individuals (**table S8**). Feather follicles were collected in the pin stage of development (an early stage of feather development shortly after emerging through the skin), placed in RNAlater at room temperature for 8 hours, and subsequently stored at -80°C until RNA extraction. Total RNA was isolated using the RNeasy Plus Mini Kit (Qiagen). Following extraction, RNA integrity was assessed using a TapeStation RNA ScreenTape (Agilent), and RNA concentration was measured using a Qubit RNA BR assay kit (ThermoFisher Scientific).

Library preparation and sequencing. Illumina strand-specific RNA-seq libraries were prepared using the TruSeq Stranded mRNA Kit, according to the manufacturer's instructions. Sequencing

of these libraries generated approximately 1,085 million paired-end reads (2 × 151 bp), with an average of ~70 million reads per individual (range: 41,911,332-97,612,226) (**table S8**).

Differential gene expression analysis. RNA-seq reads were aligned to the Indian peafowl reference genome using *HISAT2* (v.2.2.1) (*91*). Gene expression analyses were carried out using the *R* package *DESeq2* (v1.36.0) (*92*), with count data for each transcript generated with the *featureCounts* (*93*) function of the *Rsubread* package (v1.22.2) (*94*) against the annotation file. Differential expression analyses compared biological replicates of each mutant to wild-type birds (neck and wing analyzed separately). Transcripts with adjusted *P*-values (corrected for multiple testing) below 0.1 were considered significantly differentially expressed.

Allelic imbalance analyses. Allelic expression imbalance in neck and wing feather follicles was quantified in individuals heterozygous for the *black-shoulder* nonsynonymous *MC1R* mutation (see main manuscript). For each sample, we estimated the relative expression of the two alleles either by directly counting RNA-seq reads overlapping the mutation site (when coverage exceeded 50×) or by targeted amplicon sequencing when RNA-seq coverage was below that threshold. Amplicon library preparation for Illumina sequencing followed Andrade *et al.* (2019) (*95*), with the following changes: first PCR with 5'-tailed primers (insert size: 177 bp; primer forward: 5' – TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAATGAGCTCTTCCTGACG – 3'; primer reverse: 5' – GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAGCAGCATGAAGAGCATC – 3'), a touchdown PCR from 66-62°C, and PCR clean-up with 1:1 bead-to-sample volume ratio. The resulting libraries were pooled and sequenced using paired-end Illumina reads (2 × 150 bp). Reads were mapped to the reference genome using *BWA-MEM* (v0.7.17-r1188) with default parameters, and counts for the alleles were obtained by visual inspection with the *Integrative Genomics Viewer* (*IGV*) (*96*). For the neck, three heterozygous samples were analyzed ($n = 3$): two using RNA-seq and one using amplicon sequencing. For the wing, allelic proportions were obtained from two samples ($n = 2$) using amplicon sequencing only, as the RNA-seq datasets for wing follicles had <50× coverage at the mutation site.

Macroscopic photographs of feathers

The macroscopic photographs of the feathers from (i) the train eyespot region of wild-type and mutant peafowl varieties, (ii) the covert wing feathers of the common emerald dove (*Chalcophaps indica*), and (iii) the secondary wing feathers (or speculum) from the mallard duck (*Anas platyrhynchos*) were taken with a Sony Alpha 6400 combined with a Godox V1s flash. Both were installed in a setup with an angle of the flash of 13°.

Optical microscope characterization of the peafowl eyespot

Light microscope images and microscopic reflectance spectra were acquired with a correlative microscope setup consisting of a Zeiss Axio Imager Z2 optical microscope (Zeiss) combined with an MCS CCD UV-NIR spectrometer (Zeiss). The microscope was equipped with an Axiocam 820 color camera (Zeiss). Illumination was provided by a built-in halogen lamp (Osram) at 9.76 V. Images of individual barbules were acquired using a 20× objective with a numerical aperture (NA) of 0.5. Spectral measurements were performed using a 50× objective with an NA of 0.55. A silver mirror (Thorlabs) served as a 100% reflectance reference. The measured spot size was 16×16 µm. A single spectrum was recorded per spot. The integration time was adjusted to avoid detector saturation and typically ranged between 100-200 ms for most color types, and 300-400 ms for *ultramarine* samples. Internal spectrometer smoothing was set to 5, and an averaging of 10 spectra was applied by the software. Binning was set to 2 × 2. Multiple spectra of the individual barbules were recorded, and the mean and standard deviation

were calculated. For sample preparation, the individual barbules were carefully cut from the barb and adhered to a carbon pad (Plano).

Color measurement and chroma determination of the peafowl eyespot

Perceived color was quantified in the CIELAB (L*, a*, b*) color space, as defined by the International Commission on Illumination (CIE). CIELAB coordinates were calculated from the reflectance spectra recorded with the Zeiss Axio Imager Z2 microscope-spectrometer setup described above. The transformation followed standard procedures under the CIE 2006 convention, where L* represents lightness, a* corresponds to the green–red axis, and b* corresponds to the blue–yellow axis, under standard daylight illumination (D65) with a 2° observer angle and a step size of 5 nm. Chroma values were subsequently derived from these coordinates (*97*).

Spectroscopy of the peafowl eyespot

For the angle-dependent spectrometry, a customized setup was employed. A tungsten halogen lamp (HL-2000-HP-FHSA, Ocean Insight) served as the light source. The collected light was analyzed using a Flame-T-VIS-NIR-ES spectrometer (Ocean Insight). Both components were connected to collimators (74-UV-MP, Ocean Insight), which were mounted on mechanical arms and linked via glass fiber cables (105 µm, 0.22 NA Fiber Patch Cable, Thorlabs). These arms were controlled using two motorized rotational stages (Standa) to precisely adjust the angles of incidence and detection. Spectra were recorded using a custom *LabView*-based acquisition software. With this software, the integration time, averaging, and data export were controlled.

For sample preparation, the feathers were glued to a glass slide using carbon pads (Plano), ensuring a flat and stable surface for accurate measurements. Specular reflection measurements were conducted. The incidence angle was varied from 60° to 3° and subsequently from −3° to −60° in 1.5° steps. Measurements for incidence angles between 3° and −3° were not possible due to the obstruction of the light signal by the detector arm. Because angle-dependent spectra recorded from opposite sides exhibited symmetric (mirrored) peak behavior, only one side of the measurements was used for subsequent analysis (**fig. S16**). For reference measurements, a white standard (Labsphere) was used as the maximum reference, measured at the same angles as the sample. The minimum reference was obtained by recording the spectrometer signal without illumination. An integration time of 2800 ms and an averaging of 10 scans were used. For samples with a well-ordered photonic crystal (wild-type, *purple*, *ultramarine*, and *violet* feathers), the spectra were smoothed over 200 points using the Adjacent-Averaging method in *OriginPro* (v2023b) (OriginLab), and the peak position was determined based on the highest intensity. For samples exhibiting either a less well-ordered photonic crystal or a highly ordered but few-layer photonic crystal structure that resulted in low reflectance (e.g., *bronze* and *opal* feathers), the spectra were first smoothed and subsequently baseline-corrected using a spline function in *OriginPro*, as the broadband pigment background otherwise obscured the structural peak. Peak positions were then determined by applying a Gaussian fit to the peak region, also in *OriginPro*. For samples without a photonic crystal (*cameo* and *charcoal* feathers), the spectra were only smoothed (200-point), and no peak could be evaluated. Because transformations did not significantly impact the peak wavelength identified for mutants across the multiangle measurements, values of the 200-point measures were used for peak selection (**fig. S17**).

Transmission Electron Microscopy (TEM)

TEM initial sample preparation. Feathers used for TEM were collected from the same sources described for macroscopic photographs, including wild-type and mutant peafowl, the common emerald dove, and the mallard duck. Feathers were washed twice with 0.1 M phosphate buffer (PB, pH 7.4) containing 0.5% Tween 20 (Carl Roth) and stored overnight at 4°C in the same

solution. Samples were then treated for 2 hours at room temperature with a solution of 1.5% potassium ferricyanide (Merck) and 1% osmium tetroxide (>99.95%, Emsdiasum) in 0.1 M PB (pH 7.4). Afterwards, the feathers were washed three times in the same PB-Tween solution and stored overnight at 4°C.

Preparation of 0.1 M Phosphate Buffer Solution (Sørensen method). To prepare 1 liter of phosphate buffer solution (PB), 14.417 g of disodium hydrogen phosphate dihydrate ($Na_2HPO_4 \cdot 2H_2O$, ≥99%, Carl Roth) and 2.622 g of sodium dihydrogen phosphate monohydrate ($NaH_2PO_4 \cdot 1H_2O$, ≥98%, Carl Roth) were dissolved in 900 ml of Milli-Q water. The pH of the solution was adjusted to 7.4 using hydrochloric acid (HCl, ≥25%, Carl Roth) or sodium hydroxide (NaOH, ≥98%, Carl Roth). Finally, Milli-Q water was added to bring the total volume to 1000 ml.

Feather dehydration process. Feathers were progressively dehydrated at room temperature through a graded ethanol–acetone series. Samples were first immersed in 70% ethanol (≥99.8%, denatured, Carl Roth) for 60 minutes, followed by 80%, 90%, and 100% ethanol (denatured) for 30 minutes each. The dehydration continued with 100% ethanol (undenatured, ≥99.8%, Carl Roth) for 30 minutes, then a 50/50 mixture of ethanol (≥99.8%, Carl Roth) and acetone (≥99.8%, Carl Roth) for 30 minutes, followed by 100% acetone for 30 minutes. Finally, samples were immersed in 2/3 acetone and 1/3 Epon resin, and then 1/3 acetone and 2/3 Epon resin, each for 30 minutes at room temperature.

Preparation of Epon resin. Epon resin was prepared by mixing Solution A and Solution B as follows:

1) Solution A: 75 mL Glycidether 100 (Carl Roth) and 120 mL Glycidether hardener DBA (Carl Roth)
2) Solution B: 120 mL Glycidether 100 (Carl Roth) and 105 mL Glycidether hardener MNA (Carl Roth)

The two solutions were combined and supplemented with 2% accelerator DMP-30 (Carl Roth). The mixture was gently stirred until homogeneous and stored at 4°C until use.

Embedding and sectioning. Dehydrated feathers were embedded in Epon resin with glycidether accelerator DMP-30 (Carl Roth) overnight at 4°C. Polymerization was carried out by incubating samples at 60°C overnight, followed by 80°C, again overnight. The Epon blocks were cut into 40-60 nm thick sections using a Leica Ultracut UCT. These sections were treated with chloroform (≥99%, Carl Roth) to ensure a flat morphology and transferred onto copper grids coated with Pioloform FN65 (Plano) film.

Section preparation and image acquisition. Sections were stained with 10% uranyl acetate (>98%, Serva) for 10 minutes, rinsed with Milli-Q water, treated with 2.8% lead citrate (Merck) for 10 minutes, and rinsed again with Milli-Q water. Imaging was performed using a Zeiss Leo 906 transmission electron microscope (Zeiss) operated in bright-field mode at 60 kV with a TRS 2048 digital camera system.

Nanostructure measurements: Several variables related to melanosome size and spatial distribution in the keratin matrix were measured from TEM images. These variables included: number of melanosome layers (*N*), distance between layers (*a*), distance between melanosomes within a layer (*b*), cortex thickness (*c*), diameter of melanosomes (*Dm*), diameter of air pockets (*Da*), and length of melanosomes (*Lm*). Measurements were manually taken using the *straight line* and *measure* tools from *ImageJ* (*98*) software on longitudinal and transversal cuts of feather barbules. For each breed–zone combination (n = 42), 100 measurements were taken, and the mean and standard deviation were calculated.

Statistical analyses. Linear discriminant analyses (LDA) were conducted using either "phenotype" or "area" as the grouping variable. Model performance was assessed as the

percentage of individuals correctly reclassified to their original group, based on a classifier that minimized each sample's multivariate Mahalanobis distance to the group mean. Pearson's correlation coefficients (r) were calculated to identify significant associations ($P < 0.05$) between Z-score standardized averages of the original measurements and the axes of the LDA models. Analyses were implemented using the *MASS R* package (*99*). To estimate the relative contribution of each variable to total dataset variance, squared loadings were weighted, summed, and normalized across LD axes to obtain contribution percentages.

Histology of regenerating feather follicles

Regenerating feather follicles at the pin stage of development were collected from the following sources: (i) the neck region of wild-type and mutant peafowl varieties, (ii) the covert wing feathers of the common emerald dove, and (iii) the secondary wing feathers (or speculum) from the mallard duck. The samples were fixed in 4% paraformaldehyde prepared in 0.1M PBS (pH 7.4) at 4°C for four days. After fixation, the follicles were washed in PBS and stored in 70% ethanol until further processing. The follicles were then dehydrated in an ascending ethanol series, cleared in xylene, and embedded in Paraplast Plus (Sigma). Longitudinal 10 µm thick sections were cut with a Reichert-Jung 2040 Autocut Rotary Microtome. The ribbon of sections was floated on the surface of distilled water at 37°C and mounted onto gelatin-coated glass slides. The sections were heated at 65°C for 1 hour to melt the paraffin, deparaffinized in xylene, and mounted with DPX mounting medium (Sigma). Images were captured using a Zeiss Axio Imager M2 epifluorescent microscope equipped with AxioCam MRc5 camera (Zeiss) and *AxioVision Se64* (vRel.4.9.1) software (Zeiss). A 63× Plan-apochromat oil immersion objective was used for imaging.

Chemical characterization of melanin

Fully grown feathers from the same sources used in the histological analysis (see previous section) were collected for the quantification of melanin content. Feathers were kept in black plastic zip-lock bags at room temperature, protected from direct light to prevent degradation. Approximately 15 mg of each feather sample was homogenized in water using a Ten-Broeck glass homogenizer at a concentration of 10 mg/ml. An aliquot of 100 µl (equivalent to 1 mg of sample) was mixed with 900 µL of Soluene-350 (Perkin-Elmer) in a 10 ml screw-capped conical test tube (*100*). The mixture was vortexed and heated at 100˚C in a boiling water bath for 15 minutes, then cooled. This process was repeated, followed by vortexing and centrifugation at 4,000 g for 3–5 minutes. The resulting supernatant was analyzed using a spectrophotometer at 500 nm (A500) and 650 nm (A650). A blank reference sample (100 µl water and 900 µl Soluene-350) was processed under the same conditions. To analyze pyrrole-2,3,5-tricarboxylic acid (PTCA), pyrrole-2,3-dicarboxylic acid (PDCA), and thiazole-2,4,5-tricarboxylic acid (TTCA), alkaline hydrogen peroxide oxidation (AHPO) was employed (*101*), while 4-amino-3-hydroxyphenylalanine (4-AHP) and 3-amino-4-hydroxyphenylalanine (3-AHP) were quantified via hydroiodic acid (HI) hydrolysis (*102*). Absorbances due to proteins (0.021 and 0.001 per mg of feather) were subtracted from A500 and A650 values. Total melanin, eumelanin, benzothiazine-pheomelanin (BT), and benzothiazole-pheomelanin (BZ) contents were calculated by multiplying the concentrations of A500, PTCA, 4-AHP, and TTCA by factors of 101 (*100*), 80 (*101*), 9 (*103*), and 34 (*104*), respectively. The quantification of different metabolites from these pigment analyses is available in **tables S6 and S9**.

Single-cell RNA-sequencing

Tissue preparation and production of libraries for scRNA-seq. The experiments were conducted on regenerating feathers of one male emerald dove and two male mallards, taking advantage of the fact that these species express within some feathers both iridescent and non-iridescent colors

(**Fig. 6A; fig. S10**). For each species, two scRNA-seq libraries were generated per individual feather to capture the structural coloration asymmetry, with one library representing the left side and the other the right side (emerald dove: n = 4 libraries; mallard: n = 4 libraries). For the emerald dove, wing covert feathers belonging to a row of feathers displaying color asymmetry relative to the central axis of the feather itself (i.e., the rachis) were analyzed. The feathers were sampled approximately nine days after being plucked, at a time when color differences between the two sides had just started to be visible. For the mallard duck, feathers of the speculum, a bright iridescent patch of inner secondary flight feathers, were sampled. Mallard feathers were sampled 20 days after plucking during the formation of the iridescent patch.

For all the experiments, regenerating feathers were plucked, immediately placed on ice in Hanks' Balanced Salt Solution (HBSS, Gibco), and dissected under a stereomicroscope. The distal-most half of the feather, which is fully keratinized at the time of sampling, was removed (**fig. S10**). The remaining feather was incised longitudinally along the rachis and flattened by pinning its corners in HBSS. This exposed the inner pulp, which was then carefully removed with forceps. At this stage, in both emerald dove and mallard feathers, the color asymmetry is evident as a darkening of the tissue where structural coloration would emerge during later keratinocyte maturation. After confirming this difference, the feather was further cut at about 4 mm from its base, at a level where barb ridges are formed, and the proximal-most portion of the sample was kept. The left and right halves of the feather were then separated by cutting along an evident midline devoid of keratinocytes, and processed individually for cell dissociation.

The samples were dissociated in 2mg/ml dispase solution in DMEM (Gibco) for 60 min at 37°C at 300 rpm. After 10 minutes of incubation, 30 µl of Liberase solution (Roche) in DMEM (10 mg/ml) was added, and the mixture was passed through a pipette tip several times every 10 minutes. The partially dissociated tissues were pelleted by centrifugation at 400 × g at 4°C and incubated in 0.05% Trypsin/EDTA (Gibco) for 10 minutes at 37°C and 300 rpm mixing, and occasionally passed through a pipette tip. The digestion was arrested by adding 10% FBS (Gibco) in DMEM, the feather sheath was removed with forceps, and the mixture was treated with DNase I (Roche) at 37°C for 5 min and 300 rpm. The cells were then washed twice in 0.4% BSA (Sigma-Aldrich) in HBSS (with calcium and magnesium), filtered twice through a 40 µm-mesh filter (Falcon) and once through a Flowmi 40 µm Cell Strainer (Bel-Art SP Scienceware), resuspended in 0.4% BSA in HBSS, and counted in a hemocytometer after trypan blue staining. The resuspended cells were partitioned and barcoded separately using a 10x Genomics Chromium instrument (10x Genomics) following the manufacturer's protocol (Chromium Next GEM Single Cell 3′ Reagent Kits v3.1 Dual Index, Rev C) and targeting the recovery of 10,000 cells per reaction. Following quantification and quality control (LightCycle qPCR, Agilent TapeStation), the libraries were sequenced on an Illumina instrument and demultiplexed using the *CellRanger Fastq* pipeline (10x Genomics). Summary statistics for the single-cell libraries are given in **table S10**.

Preprocessing of the scRNA-seq data. The scRNA-seq libraries were pre-processed individually using the *count* pipeline in *Cell Ranger* (v7.0.1; 10x Genomics). Briefly, the reads were aligned to the reference genomes (emerald dove: this study and described below; mallard: GCF_047663525.1) using the splicing-aware aligner *STAR* (*65*) as implemented in *Cell Ranger*. Uniquely mapped reads within the start and end coordinates of each gene were considered for UMI counts, and cells were distinguished from empty droplets using an algorithm based on the *EmptyDrop* method (*105*). Filtered count matrices were imported into the *R* package *Seurat* (v5.3.0) (*106*) and further filtered to remove low-quality cells. For the emerald dove (n = 4) and mallard (n = 4) libraries, cells with UMI counts between 2,000 and 18,000 and between 2,000 and 20,000, respectively, and expressing more than 750 features were retained. After filtering, 10,213 emerald dove cells and 20,991 mallard cells remained. Gene expression matrix processing (data normalization and scaling) and dimensionality reduction by Principal Components Analysis

(PCA) were performed independently for each library in *Seurat* with default settings. For each species separately, the libraries were integrated using the *HarmonyIntegration* method as implemented in the *IntegrateLayers* function in *Seurat*.

Cluster annotation. Unsupervised clustering of cells was performed on the integrated data using the *FindNeighbors* (*k.param* = 50, $dims_{dove}$ = 50; $dims_{mallard}$ = 30) and the *FindClusters* ($resolution_{dove}$ = 0.14; $resolution_{mallard}$ = 0.05) functions in *Seurat.* Differential gene expression analyses between each cluster and all remaining cells in the dataset, as well as between selected cluster pairs, were conducted with the Wilcoxon rank-sum test as implemented in the *FindAllMarkers* and *FindMarkers* functions in *Seurat,* respectively. Cluster annotation was manually curated based on the expression of known marker genes and the top differentially expressed genes in each cluster. An overview of the cluster annotations and representative marker genes is shown in **figs S11-S14**, and complete lists of the top differentially expressed genes for each cluster are provided in **Data S3** and S**4**.

Within each dataset, clusters of cells showing enriched expression of keratin genes (e.g., *Keratin-14*, *Keratin-19*, *Keratin-8*, *Keratin-5*) and known keratinocyte markers (e.g., *SHH* and *COL17A1*) (*107–109*) were collectively labeled as keratinocytes. In both datasets, two distinct clusters of cells showed enriched expression of melanocyte-specific markers such as *TYR*, *TYRP1*, *PMEL*, and *MLANA*. Within each dataset, pairwise differential expression analyses between these clusters identified one population of cells (labeled as melanocytes 1) showing upregulation of several melanocyte markers (e.g., *TRPM1*: Log_2FC_{dove} = 4.6; $P < 0.0001$; $Log_2FC_{mallard}$ = 3.7, $P < 0.0001$). In both datasets, the second population (labeled as melanocytes 2) showed enriched expression of several glutathione S-transferases (e.g., Hematopoietic Prostaglandin D Synthase*, HPGDS*: Log_2FC_{dove} = 1.76; $P < 0.0001$; $Log_2FC_{mallard}$ = 1.29, $P < 0.0001$). A previous study in quails identified two distinct melanocyte lineages committed to the production of eumelanin or pheomelanin within embryonic skin tracts, eventually forming black or yellow feather stripes (*110*). These populations were negative and positive for the melanoblast/melanocyte early marker (MelEM) monoclonal antibody, which binds a glutathione S-transferase alpha class subunit and labels melanoblasts and early melanocytes in the avian embryo (*111*, *112*). Increased glutathione S-transferase expression and decreased melanocyte-specific marker gene expression suggest that these cells are early differentiating melanocytes, possibly progressing towards pheomelanin production. Based on the common expression of many melanocyte markers, these populations of cells were collectively labeled as melanocytes and were considered together for some analyses.

Additional small but clearly distinct clusters were manually annotated. In both datasets, a small cluster of cells expressing erythrocyte markers, such as *HBAA* and *HBB*, was manually assigned and annotated as erythrocytes (emerald dove: n = 30 cells; mallard: n = 67 cells). In the mallard dataset, co-expression of erythrocyte markers and some keratin genes suggests the contamination of small numbers of low-quality keratinocyte cells within the erythrocyte cluster. In the emerald dove dataset, one small cluster of cells (n = 29 cells) showed enriched expression of neural crest-derived cell genes (e.g. *FOXD3* and *NGFR*) and likely represents the melanoblast population at the base of the feather follicle. In the absence of canonical melanoblast markers (e.g., *DCT, EDNRB2*), this cluster was manually assigned and labeled as neural crest-derived cells. Overall, these changes represented minor manual adjustments to the clustering output, limited to clearly distinct and well-characterized cell populations, and did not affect the overall interpretation of the data. The remaining clusters in both datasets were left unchanged from the clustering provided by *Seurat* and were annotated based on the expression of marker genes for cell types previously described in regenerating feather follicles (*7*).

For the emerald dove, the final annotation of cells consisted in 11 clusters: (1) keratinocytes, n = 4,261; (2) mesenchymal cells 1, n = 1,716; (3) melanocytes 1, n = 1,430; (4) endothelial cells, n = 1,215; (5) melanocytes 2, n = 566; (6) leukocytes 1, n = 373; (7)

mesenchymal cells 2, n = 276; (8) leukocytes 2, n = 239; (9) leukocytes 3, n = 77; (10) erythrocytes, n = 31, and (11) neural crest-derived cells, n = 29. For the mallard, the final annotation consisted of 8 clusters: (1) keratinocytes, n = 14,312; (2) mesenchymal cells 1, n = 4,376; (3) melanocytes 1, n = 855; (4) leukocytes, n = 491; (5) endothelial cells, n = 406; (6) melanocytes 2, n = 334; (7) mesenchymal cells 2, n = 150; and (8) erythrocytes, n = 67.

Tissue composition analyses. For both datasets, cell counts per cluster were extracted from each library (n = 4 libraries per species) and normalized by library total cell counts. Cell-type proportions in the samples were compared between conditions (iridescent color vs. non-iridescent color) with a Welch Two Sample t-test in *R* (n = 2 replicates per condition). Cell-type counts and summary test statistics are reported in **table S11**.

Differential gene expression analyses. For each species and for each cell type, differential expression analyses were performed using a pseudo-bulk approach. For all the analyses, differentially expressed genes with $Log_2FC \geq 0.5$ and P-value ≤ 0.05 were considered. To test for differences between feather samples developing iridescent or non-iridescent colors (i.e., color types), summed UMI counts within each sample ($n_{iridescent}$ = 2 samples per dataset, $n_{non\text{-}iridescent}$ = 2 samples per dataset) were used as inputs for *DESeq2* (v1.48.1) with default parameters as implemented in the *FindMarkers* function in *Seurat*. Lists of differentially expressed genes per cell type for the emerald dove and the mallard are reported in **Data S1 and S2**, respectively.

Gene ontology enrichment analyses. For each species, gene ontology (GO) enrichment analyses were carried out for the lists of melanocyte-specific differentially expressed genes (considering the merged populations of melanocytes 1 and 2) using the *topGO* package (v.2.59.0) (https://bioconductor.org/packages/topGO) in *R*. For each dataset, enrichment was assessed by comparing the frequencies of GO term annotations in the differentially expressed gene sets against background GO term frequencies in the species' genomes, followed by statistical testing by Fisher's exact test. GO terms associated with the differentially expressed gene sets for the emerald dove and the mallard are reported in **Data S5 and S6**, respectively.

Emerald dove genome assembly and annotation. The individual used to produce the draft reference genome sequence of the emerald dove was one adult male obtained from a licensed breeder in Portugal. Whole blood was collected using a sterile needle into a heparin-free capillary, transferred into a 10ml Vacutainer K2 EDTA tube (BD Biosciences), and immediately snap-frozen in liquid nitrogen. The sample was subsequently shipped to the genomic services provider Macrogen (Seoul, South Korea) in dry ice. High molecular weight DNA was extracted from 140 µl of blood using the Wizard HMW DNA Extraction kit (Promega). DNA quantity and integrity were assessed using a Victor Nivo Multimode Microplate Reader (PerkinElmer) with the QuantiFluor® dsDNA System (Promega) method, the Femto Pulse System (Agilent), and a N120 NanoPhotometer (IMPLEN). The average fragment length of the extracted DNA was estimated to be approximately 44 kb.

A Pacific Biosciences (PacBio) Revio sequencing library was prepared from 7 µg of DNA input. gDNA was sheared with the Megaruptor® 3 (Diagenode) and both purified and size-selected using AMPure PB magnetic beads (Pacific Biosciences). The size distribution was determined using the Femto Pulse System (Agilent) for all size quality controls. The library insert size was within the optimal size range. A total of 10 µL of library was prepared using the PacBio SMRTbell prep kit 3.0 and SMRTbell templates were annealed using the Revio polymerase kit (Pacific Biosciences). Sequencing was performed using a Revio sequencing plate and Revio SMRT Cell tray. Each SMRT cell was captured using a 24-hour movie time using the PacBio Revio (Pacific Biosciences) sequencing platform by Macrogen (Seoul, South Korea). The subsequent steps are based on the PacBio Sample Net-Shared Protocol, available at https://www.pacb.com/.

A total of 6,372,772 HiFi long reads with an average length of 13,480 bp were generated, which corresponds to ~71× coverage of the approximate size of a bird genome (~1.2 Gb). Contig assembly was performed using *Hifiasm* (v0.20.0-r639) (*113*), resulting in 1,283,043,043 bp of primary assembly (contigs $n = 315$) with a contig N50 of 33 Mb and a maximum length of 121.1 Mb. These contigs were scaffolded into a pseudochromosome assembly using *RagTag* (v2.1.0) (*114*), with the rock pigeon *bColLiv1.pat.W.v2* genome assembly (GCF_036013475.1) serving as the reference. The final assembly comprised 219 scaffolds with an N50 of 83.8 Mb and a maximum scaffold length of 215.1 Mb. Completeness analysis of the assembly was performed using *BUSCO* (v5.8.2) and the 8,838 single-copy avian ortholog genes (*aves_odb10*), which indicates that our genome assembly was 97.1% complete. Annotation of the draft reference genome was conducted using *Liftoff* (v1.6.3) (*115*) by lifting over gene models from the annotation available for the rock pigeon genome. Genes of interest for the single-cell analyses were manually inspected.

DATA AVAILABILITY

The whole-genome sequencing (Illumina), RNA sequencing, and single-cell RNA sequencing data generated in this study are available at the NCBI BioProject under accession number PRJNA1216069. The reference genome assembly and associated data, including Oxford Nanopore sequencing, Hi-C, and linked-read sequencing, are available under the ENA project accession PRJEB87866. The raw data supporting the nanostructural and optical analyses are available in the public repository Zenodo at the following DOI: XXXXXX.

ACKNOWLEDGEMENTS

We thank the numerous breeders who kindly provided feather and tissue samples for data analyses. We are also grateful to Jens Poulsen and Clucking Fens Farm (www.cluckingfens.co.uk) for peacock photographs used in Fig. 1. We thank Ibanidis Lda (Versele-Laga distributor), ZooService Lda (Manitoba distributor), and Papa d'Ovo (Domus Molinari distributor) as sponsoring partners of food, husbandry cages, and accessories. We thank Joana C. Carvalho for helping produce the main figures. We thank Nico Nees for kindly providing the code used for the chroma calculations, Bettina Winzer for her assistance in aligning the angle-dependent spectrometer, and Carmen Rubach for helpful discussions. This work was supported by the European Research Council under the European Union's Horizon 2020 research and innovation program to MC (grant agreement No. 101000504); by the Portuguese Foundation for Science and Technology (FCT, https://www.fct.pt) research contracts to MC (CEECINST/00014/2018/CP1512/CT0002), PA (2020.01405.CEECIND/CP1601/CT0011), and PMA (2020.01494.CEECIND); by research fellowships to CIM (SFRH/BD/147030/2019) and RAf (UI/BD/154451/2022), in the scope of the Biodiversity, Genetics and Evolution (BIODIV) PhD program; by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 416229255 – SFB 1411; by an EASI-Genomics project under CNAG-CRG coordination, funded by the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 824110); by national funds through FCT to CE3C (UIDB/00329/2020, https://doi.org/10.54499/UIDB/00329/2020); and by funding from Vetenskapsrådet (VR 2017-02907), and the Knut and Alice Wallenberg Foundation (KAW 2023.0160 and KAW 2024.0058) to LA.

AUTHOR CONTRIBUTIONS

Conceptualization: M.C., N.V., J.C.C., L.A., S.B., C.B., R.A., P.M.A.
Methodology: S.B., P.M.A., P.P., P.A., R.Af., S.Af., J.Ca., R.J.L., J.M.A., M.F.A., R.A., C.B., M.P.J.N., J.B., A.Hi., A.He., E.S., B.A.Z., C.I.M., C.F., M.A.A., F.C., J.G.G., C.Z., M.G., T.A., S.I., K.W.

Investigation: S.B., P.M.A., P.P., P.A., R.Af., S.Af., J.Ca., R.J.L., J.M.A., M.F.A., R.A., C.B., M.P.J.N., J.B., A.Hi., A.He., E.S., B.A.Z., C.I.M., C.F., M.A.A., F.C., J.G.G., C.Z., M.G., T.A., S.I., K.W.
Data curation: S.B., P.M.A., P.P., P.A., R.A., C.B., J.B., A.Hi., A.He., E.S., B.A.Z., C.F., F.C., J.G.G., C.Z., M.G., T.A.
Formal analysis: S.B., M.F.A., R.A., P.P., C.B., M.P.J.N., J.B., A.Hi., A.He., E.S., B.A.Z.
Visualization: S.B., R.A., C.B., C.F., J.A.A., T.A., M.P.J.N., J.B., A.Hi., A.He., E.S., B.A.Z.
Supervision: M.C., N.V.

DECLARATION OF INTERESTS

S.B., R.A., P.M.A. and M.C. are inventors on a pending patent application related to the technology described in this work. S.B., P.M.A. and M.C. are founders and shareholders of *Animal Genotyping Innovations (AGI), Lda*. The remaining authors declare no competing interests.

REFERENCES

1. S. Kinoshita, Helen Ghiradella, Lars Olof Björn, "Spectral Tuning in Biology II: Structural Color" in *Photobiology*, Björn L., Ed. (Springer, New York, 2015).
2. I. C. Cuthill, W. L. Allen, K. Arbuckle, B. Caspers, G. Chaplin, M. E. Hauber, G. E. Hill, N. G. Jablonski, C. D. Jiggins, A. Kelber, J. Mappes, J. Marshall, R. Merrill, D. Osorio, R. Prum, N. W. Roberts, A. Roulin, H. M. Rowland, T. N. Sherratt, J. Skelhorn, M. P. Speed, M. Stevens, M. C. Stoddard, D. Stuart-Fox, L. Talas, E. Tibbetts, T. Caro, The biology of color. *Science* **357**, eaan0221 (2017).
3. M. D. Shawkey, L. D'Alba, Interactions between colour-producing mechanisms and their effects on the integumentary colour palette. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160536 (2017).
4. P. Vukusic, J. R. Sambles, Photonic structures in biology. *Nature* **424**, 852–855 (2003).
5. G. E. Hill, K. J. McGraw, *Bird Coloration. In Mechanisms and Measurements, 1* (Harvard University Press, 2006).
6. T. F. Cooke, C. R. Fischer, P. Wu, T. X. Jiang, K. T. Xie, J. Kuo, E. Doctorov, A. Zehnder, C. Khosla, C. M. Chuong, C. D. Bustamante, Genetic mapping and biochemical basis of yellow feather pigmentation in budgerigars. *Cell* **171**, 427-439 (2017).
7. R. Arbore, S. Barbosa, J. Brejcha, Y. Ogawa, Y. Liu, M. P. J. Nicolaï, P. Pereira, S. J. Sabatino, A. Cloutier, E. S. K. Poon, C. I. Marques, P. Andrade, G. Debruyn, S. Afonso, R. Afonso, S. G. Roy, U. Abdu, R. J. Lopes, P. Mojzeš, P. Marík, S. Y. W. Sin, M. A. White, P. M. Araújo, J. C. Corbo, M. Carneiro, A molecular mechanism for bright color variation in parrots. *Science* **386**, eadp7710 (2024).
8. M. A. Gazda, P. M. Araújo, R. J. Lopes, M. B. Toomey, P. Andrade, S. Afonso, C. Marques, L. Nunes, P. Pereira, S. Trigo, G. E. Hill, J. C. Corbo, M. Carneiro, A genetic mechanism for sexual dichromatism in birds. *Science* **368**, 1270–1274 (2020).
9. R. Price-Waldman, M. C. Stoddard, Avian Coloration Genetics: Recent Advances and Emerging Questions. *J. Hered.* **112**, 395–416 (2021).
10. R. O. Prum, "Anatomy, physics, and evolution of avian structural colors" in *Bird Coloration, Volume 1: Mechanisms and Measurements*, G. E. Hill, K. J. McGraw, Eds. (Harvard University Press, 2006).
11. V. Saranathan, C. Finet, Cellular and developmental basis of avian structural coloration. *Curr. Opin. Genet. Dev.* **69**, 56–64 (2021).
12. M. C. Stoddard, R. O. Prum, How colorful are birds? Evolution of the avian plumage color gamut. *Behav. Ecol.* **22**, 1042–1052 (2011).
13. M. F. Land, The physics and biology of animal reflectors. Prog. Biophys. Mol. Bio. **24**, 75-106 (1972).
14. C. M. Eliason, R. Maia, M. D. Shawkey, Modular color evolution facilitated by a complex nanostructure in birds. *Evolution* **69**, 357–367 (2015).

15. R. Maia, D. R. Rubenstein, M. D. Shawkey, Key ornamental innovations facilitate diversification in an avian radiation. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10687–10692 (2013).
16. M. P. J. Nicolaï, B. Van Hecke, S. Rogalla, G. Debruyn, R. C. K. Bowie, N. J. Matzke, S. J. Hackett, L. D'Alba, M. D. Shawkey, The evolution of multiple color mechanisms is correlated with diversification in sunbirds (Nectariniidae). *Syst Biol* **73**, 343–354 (2024).
17. K. K. Nordén, J. W. Faber, F. Babarović, T. L. Stubbs, T. Selly, J. D. Schiffbauer, P. Peharec Štefanić, G. Mayr, F. M. Smithwick, J. Vinther, Melanosome diversity and convergence in the evolution of iridescent avian feathers-Implications for paleocolor reconstruction. *Evolution* **73**, 15–27 (2019).
18. C. M. Eliason, M. P. J. Nicolaï, C. Bom, E. Blom, L. D'Alba, M. D. Shawkey, Transitions between colour mechanisms affect speciation dynamics and range distributions of birds. *Nat. Ecol. Evol.* **8**, 1723–1734 (2024).
19. M. P. J. Nicolaï, G. Debruyn, R. Bowie, S. Edwards, M. Carneiro, L. D'Alba, M. Shawkey, The evolution of brilliant iridescence in birds. [Preprint] (2025). https://doi.org/10.21203/rs.3.rs-8020267/v1.
20. R. Dakin, R. Montgomerie, Eye for an eyespot: how iridescent plumage ocelli influence peacock mating success. *Behav. Ecol.* **24**, 1048–1057 (2013).
21. S. M. Doucet, M. G. Meadows, Iridescence: a functional perspective. *J. R. Soc. Interface* **6**, S115–132 (2009).
22. V. P. Patil, J. D. Sandt, M. Kolle, J. Dunkel, Topological mechanics of knots and tangles. *Science* **367**, 71–75 (2020).
23. B. E. Droguet, H.-L. Liang, B. Frka-Petesic, R. M. Parker, M. F. L. De Volder, J. J. Baumberg, S. Vignolini, Large-scale fabrication of structurally coloured cellulose nanocrystal films and effect pigments. *Nat. Mater.* **21**, 352–358 (2022).
24. N. Vogel, S. Utech, G. T. England, T. Shirman, K. R. Phillips, N. Koay, I. B. Burgess, M. Kolle, D. A. Weitz, J. Aizenberg, Color from hierarchy: Diverse optical properties of micron-sized spherical colloidal assemblies. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10845–10850 (2015).
25. K. K. Nordén, C. M. Eliason, M. C. Stoddard, Evolution of brilliant iridescent feather nanostructures. *Elife* **10**, e71179 (2021).
26. R. Hooke, *Micrographia, or, Some Physiological Descriptions of Minute Bodies Made by Magnifying Glasses with Observations and Inquiries Thereupon* (1665).
27. I. Newton, G. W. Hemming, *Opticks: Or, A Treatise of the Reflections, Refractions, Inflexions and Colours of Light : Also Two Treatises of the Species and Magnitude of Curvilinear Figures* (Printed for Sam. Smith, and Benj. Walford, London, 1704).
28. D.-J. Jeon, S. Paik, S. Ji, J.-S. Yeo, Melanin-based structural coloration of birds and its biomimetic applications. *Appl. Microsc.* **51**, 14 (2021).
29. H. Durrer, Schillerfarben der Vogelfeder als Evolutionsproblem. *Denkschr. Schweiz Naturforsch. Ges.* **14**, 1–126 (1977).
30. J. Zi, X. Yu, Y. Li, X. Hu, C. Xu, X. Wang, X. Liu, R. Fu, Coloration strategies in peacock feathers. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12576–12578 (2003).
31. P. Freyer, D. G. Stavenga, Biophotonics of diversely coloured peacock tail feathers. *Faraday Discuss* **223**, 49–62 (2020).
32. T. Okazaki, Reflectance Spectra of Peacock Feathers and the Turning Angles of Melanin Rods in Barbules. *Zoolog. Sci.* **35**, 86–91 (2018).
33. N. I. Mundy, Colouration genetics: Pretty polymorphic parrots. *Current Biology* **28**, R113–R114 (2018).
34. R. Maia, L. D'Alba, M. D. Shawkey, What makes a feather shine? A nanostructural basis for glossy black colours in feathers. *Proc. R. Soc. Lond. B Biol. Sci.* **278**, 1973–1980 (2011).
35. D. R. Rubenstein, A. Corvelo, M. D. MacManes, R. Maia, G. Narzisi, A. Rousaki, P. Vandenabeele, M. D. Shawkey, J. Solomon, Feather Gene Expression Elucidates the Developmental Basis of Plumage Iridescence in African Starlings. *J. Hered.* **112**, 417–429 (2021).
36. G. Gao, M. Xu, C. Bai, Y. Yang, G. Li, J. Xu, Z. Wei, J. Min, G. Su, X. Zhou, J. Guo, Y. Hao, G. Zhang, X. Yang, X. Xu, R. B. Widelitz, C.-M. Chuong, C. Zhang, J. Yin, Y. Zuo, Comparative genomics and transcriptomics of *Chrysolophus* provide insights into the evolution of complex plumage coloration. *Gigascience* **7**, giy113 (2018).

37. H. Dürrer, Schillerfarben beim Pfau (*Pavo Cristatus* L.). *Verh. Naturforsch. Ges. Basel* **73** (1962).
38. R. Sarangarajan, R. E. Boissy, Tyrp1 and oculocutaneous albinism type 3. *Pigment Cell Res.* **14**, 437–444 (2001).
39. J. Li, B. Bed'hom, S. Marthey, M. Valade, A. Dureux, M. Moroldo, C. Péchoux, J. Coville, D. Gourichon, A. Vieaud, B. Dorshorst, L. Andersson, M. Tixier-Boichard, A missense mutation in *TYRP1* causes the chocolate plumage color in chicken and alters melanosome structure. *Pigment Cell Melanoma Res.* **32**, 381–390 (2019).
40. C. R. Goding, H. Arnheiter, MITF—the first 25 years. *Genes Dev.* **33**, 983–1007 (2019).
41. N. I. Mundy, N. S. Badcock, T. Hart, K. Scribner, K. Janssen, N. J. Nadeau, Conserved genetic basis of a quantitative plumage trait involved in mate choice. *Science* **303**, 1870–1873 (2004).
42. D. Schwochow, S. Bornelöv, T. Jiang, J. Li, D. Gourichon, B. Bed'Hom, B. J. Dorshorst, C. Chuong, M. Tixier-Boichard, L. Andersson, The feather pattern *autosomal barring* in chicken is strongly associated with segregation at the *MC1R* locus. *Pigment Cell Melanoma Res.* **34**, 1015–1028 (2021).
43. O. Vidal, J. Viñas, C. Pla, Variability of the melanocortin 1 receptor (*MC1R*) gene explains the segregation of the bronze locus in turkey (*Meleagris gallopavo*). *Poult. Sci.* **89**, 1599–1602 (2010).
44. H. E. Hoekstra, Genetics, development and evolution of adaptive pigmentation in vertebrates. *Hered.* **97**, 222–234 (2006).
45. L. Weiner, W. Fu, W. J. Chirico, J. L. Brissette, Skin as a living coloring book: how epithelial cells create patterns of pigmentation. *Pigment Cell Melanoma Res.* **27**, 1014–1031 (2014).
46. T. Hirobe, Role of dermal factors involved in regulating the melanin and melanogenesis of mammalian melanocytes in normal and abnormal skin. *Int. J. Mol. Sci.* **25**, 4560 (2024).
47. R. Maia, R. H. F. Macedo, M. D. Shawkey, Nanostructural self-assembly of iridescent feather barbules through depletion attraction of melanosomes during keratinization. *J. R. Soc. Interface* **9**, 734–743 (2012).
48. M. Giraldo, J. Sosa, D. Stavenga, Feather iridescence of *Coeligena* hummingbird species varies due to differently organized barbs and barbules. *Biol. Lett.* **17**, 20210190 (2021).
49. Y. Liu, L. Hong, K. Wakamatsu, S. Ito, B. Adhyaru, C. Cheng, C. R. Bowers, J. D. Simon, Comparison of structural and chemical properties of black and red human hair melanosomes. *Photochem. Photobiol.* **81**, 135–144 (2005).
50. J. A. Brumbaugh, Ultrastructural differences between forming eumelanin and pheomelanin as revealed by the pink-eye mutation in the fowl. *Dev. Biol.* **18**, 375–390 (1968).
51. E. D. Enbody, C. G. Sprehn, A. Abzhanov, H. Bi, M. P. Dobreva, O. G. Osborne, C.-J. Rubin, P. R. Grant, B. R. Grant, L. Andersson, A multispecies *BCO2* beak color polymorphism in the Darwin's finch radiation. *Curr. Biol.* **31**, 5597-5604.e7 (2021).
52. D. E. Wood, S. L. Salzberg, Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* **15**, R46 (2014).
53. S. Marco-Sola, M. Sammeth, R. Guigó, P. Ribeca, The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat. Methods* **9**, 1185–1188 (2012).
54. N. C. Durand, M. S. Shamim, I. Machol, S. S. P. Rao, M. H. Huntley, E. S. Lander, E. L. Aiden, Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
55. A. V. Zimin, D. Puiu, M.-C. Luo, T. Zhu, S. Koren, G. Marçais, J. A. Yorke, J. Dvořák, S. L. Salzberg, Hybrid assembly of the large and highly repetitive genome of *Aegilops tauschii*, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm. *Genome Res.* **27**, 787–792 (2017).
56. A. V. Zimin, G. Marçais, D. Puiu, M. Roberts, S. L. Salzberg, J. A. Yorke, The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677 (2013).
57. M. Kolmogorov, J. Yuan, Y. Lin, P. A. Pevzner, Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
58. S. D. Jackman, L. Coombe, J. Chu, R. L. Warren, B. P. Vandervalk, S. Yeo, Z. Xue, H. Mohamadi, J. Bohlmann, S. J. M. Jones, I. Birol, Tigmint: correcting assembly errors using linked reads from large molecules. *BMC Bioinformatics* **19**, 393 (2018).

59. L. Coombe, J. Zhang, B. P. Vandervalk, J. Chu, S. D. Jackman, I. Birol, R. L. Warren, ARKS: chromosome-scale scaffolding of human genome drafts with linked read kmers. *BMC Bioinformatics* **19**, 234 (2018).
60. R. L. Warren, C. Yang, B. P. Vandervalk, B. Behsaz, A. Lagman, S. J. M. Jones, I. Birol, LINKS: Scalable, alignment-free scaffolding of draft genomes with long reads. *Gigascience* **4**, 35 (2015).
61. O. Dudchenko, S. S. Batra, A. D. Omer, S. K. Nyquist, M. Hoeger, N. C. Durand, M. S. Shamim, I. Machol, E. S. Lander, A. P. Aiden, E. L. Aiden, De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).
62. O. Dudchenko, M. S. Shamim, S. S. Batra, N. C. Durand, N. T. Musial, R. Mostofa, M. Pham, B. Glenn St Hilaire, W. Yao, E. Stamenova, M. Hoeger, S. K. Nyquist, V. Korchina, K. Pletch, J. P. Flanagan, A. Tomaszewicz, D. McAloose, C. Pérez Estrada, B. J. Novak, A. D. Omer, E. L. Aiden, The Juicebox Assembly Tools module facilitates *de novo* assembly of mammalian genomes with chromosome-length scaffolds for under $1000. bioRxiv 254797 (2018). https://doi.org/10.1101/254797.
63. G. Marçais, A. L. Delcher, A. M. Phillippy, R. Coston, S. L. Salzberg, A. Zimin, MUMmer4: A fast and versatile genome alignment system. *PLoS Comput. Biol.* **14**, e1005944 (2018).
64. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
65. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
66. M. Pertea, G. M. Pertea, C. M. Antonescu, T.-C. Chang, J. T. Mendell, S. L. Salzberg, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
67. Y. S. Niknafs, B. Pandian, H. K. Iyer, A. M. Chinnaiyan, M. K. Iyer, TACO produces robust multisample transcriptome assemblies from RNA-seq. *Nat. Methods* **14**, 68–70 (2017).
68. D. Mapleson, L. Venturini, G. Kaithakottil, D. Swarbreck, Efficient and accurate detection of splice junctions from RNA-seq with Portcullis. *Gigascience* **7**, giy131 (2018).
69. B. J. Haas, Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654–5666 (2003).
70. O. Gotoh, Direct mapping and alignment of protein sequences onto genomic sequence. *Bioinformatics* **24**, 2438–2444 (2008).
71. M. Stanke, S. Waack, Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, ii215–ii225 (2003).
72. A. Lomsadze, V. Ter-Hovhannisyan, Y. O. Chernoff, M. Borodovsky, Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* **33**, 6494–6506 (2005).
73. B. J. Haas, S. L. Salzberg, W. Zhu, M. Pertea, J. E. Allen, J. Orvis, O. White, C. R. Buell, J. R. Wortman, Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* **9**, R7 (2008).
74. A. Conesa, S. Götz, J. M. García-Gómez, J. Terol, M. Talón, M. Robles, Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).
75. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
76. P. Jones, D. Binns, H.-Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S.-Y. Yong, R. Lopez, S. Hunter, InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
77. X. Cui, Z. Lu, S. Wang, J. Jing-Yan Wang, X. Gao, CMsearch: simultaneous exploration of protein sequence space and structure space improves not only protein homology detection but also protein structure prediction. *Bioinformatics* **32**, i332–i340 (2016).
78. E. P. Nawrocki, S. R. Eddy, Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933–2935 (2013).

79. E. P. Nawrocki, S. W. Burge, A. Bateman, J. Daub, R. Y. Eberhardt, S. R. Eddy, E. W. Floden, P. P. Gardner, T. A. Jones, J. Tate, R. D. Finn, Rfam 12.0: Updates to the RNA families database. *Nucleic Acids Res* **43**, D130-137 (2015).
80. T. M. Lowe, S. R. Eddy, tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
81. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
82. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
83. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297-1303 (2010).
84. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
85. P. Cingolani, A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, D. M. Ruden, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **6**, 80-92 (2014).
86. B. L. Browning, Y. Zhou, S. R. Browning, A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* **103**, 338–348 (2018).
87. B. L. Browning, X. Tian, Y. Zhou, S. R. Browning, Fast two-stage phasing of large-scale sequence data. *The Am. J. Hum. Genet.* **108**, 1880–1890 (2021).
88. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012).
89. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, P. C. Sham, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
90. S. D. Turner, qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* **3**, 731 (2018).
91. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
92. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
93. Y. Liao, G. K. Smyth, W. Shi, FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
94. Y. Liao, G. K. Smyth, W. Shi, The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **47**, e47–e47 (2019).
95. P. Andrade, C. Pinho, G. Perez i de Lanuza, S. Afonso, J. Brejcha, C.-J. Rubin, O. Wallerman, P. Pereira, S. J. Sabatino, A. Bellati, D. Pellitteri-Rosa, Z. Bosakova, M. A. Carretero, N. Feiner, P. Marsik, F. Pauperio, D. Salvi, L. Soler, G. M. While, T. Uller, E. Font, L. Andersson, M. Carneiro, Regulatory changes in pterin and carotenoid genes underlie balanced color polymorphisms in the wall lizard. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5633–5642 (2019).
96. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24-26 (2011).
97. Commission Internationale de l'Éclairage (CIE). Colorimetry 3rd edn, CIE Publication 15:2004 (Wiley–IS&T Series in Imaging Science and Technology, Wiley, 2007).
98. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
99. B. Ripley, B. Venables. MASS: support functions and datasets for Venables and Ripley's MASS. R package version 7.3-51 (2009).
100. H. Ozeki, S. Ito, K. Wakamatsu, A. J. Thody, Spectrophotometric characterization of eumelanin and pheomelanin in hair. *Pigment Cell Res.* **9**, 265–270 (1996).
101. S. Ito, Y. Nakanishi, R. K. Valenzuela, M. H. Brilliant, L. Kolbe, K. Wakamatsu, Usefulness of alkaline hydrogen peroxide oxidation to analyze eumelanin and pheomelanin in various tissue

samples: application to chemical analysis of human hair melanins. *Pigment Cell Melanoma Res.* **24**, 605–613 (2011).
102. K. Wakamatsu, S. Ito, J. L. Rees, The usefulness of 4-Amino-3-hydroxyphenylalanine as a specific marker of pheomelanin. *Pigment Cell Res.* **15**, 225–232 (2002).
103. M. d'Ischia, K. Wakamatsu, A. Napolitano, S. Briganti, J. Garcia-Borron, D. Kovacs, P. Meredith, A. Pezzella, M. Picardo, T. Sarna, J. D. Simon, S. Ito, Melanins and melanogenesis: methods, standards, protocols. *Pigment Cell Melanoma Res.* **26**, 616–633 (2013).
104. S. Del Bino, S. Ito, J. Sok, Y. Nakanishi, P. Bastien, K. Wakamatsu, F. Bernerd, Chemical analysis of constitutive pigmentation of human epidermis reveals constant eumelanin to pheomelanin ratio. *Pigment Cell Melanoma Res.* **28**, 707–717 (2015).
105. A. T. L. Lun, S. Riesenfeld, T. Andrews, T. P. Dao, T. Gomes, J. C. Marioni, EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
106. Y. Hao, S. Hao, E. Andersen-Nissen, W. M. Mauck, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).
107. E. Cohen, C. Johnson, C. J. Redmond, R. R. Nair, P. A. Coulombe, Revisiting the significance of keratin expression in complex epithelia. *J. Cell Sci.* **135**, jcs260594 (2022).
108. M. Nishimura, W. Nishie, Y. Shirafuji, S. Shinkuma, K. Natsuga, H. Nakamura, D. Sawamura, K. Iwatsuki, H. Shimizu, Extracellular cleavage of collagen XVII is essential for correct cutaneous basement membrane formation. *Hum. Mol. Genet.* **25**, 328–339 (2016).
109. M. Yu, P. Wu, R. B. Widelitz, C.-M. Chuong, The morphogenesis of feathers. *Nature* **420**, 308–312 (2002).
110. M. Inaba, C.-M. Chuong, Avian pigment pattern formation: Developmental control of macro- (across the body) and micro- (within a Feather) level of pigment patterns. *Front. Cell Dev. Biol.* **8**, 620 (2020).
111. V. Nataf, P. Mercier, C. Ziller, N. M. Le Douarin, Novel markers of melanocyte differentiation in the avian embryo. *Exp. Cell Res.* **207**, 171–182 (1993).
112. V. Nataf, P. Mercier, B. De Néchaud, J.-C. Guillemot, J. Capdevielle, F. Lapointe, N. M. Le Douarin, Melanoblast/melanocyte early marker (MelEM) is a glutathione S-transferase subunit. *Exp. Cell Res.* **218**, 394–400 (1995).
113. H. Cheng, G. T. Concepcion, X. Feng, H. Zhang, H. Li, Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
114. M. Alonge, L. Lebeigle, M. Kirsche, K. Jenike, S. Ou, S. Aganezov, X. Wang, Z. B. Lippman, M. C. Schatz, S. Soyk, Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* **23**, 258 (2022).
115. A. Shumate, S. L. Salzberg, Liftoff: accurate mapping of gene annotations. *Bioinformatics* **37**, 1639–1643 (2021).

Figure 1. Mendelian color mutants in the Indian peafowl (*Pavo cristatus*). This figure illustrates the striking phenotypic effects of Mendelian color mutations on the overall appearance and eyespot coloration in peacocks. From left to right: wild-type, *ultramarine*, *purple*, *violet*, *opal*, *bronze*, *cameo*, *charcoal*, and *white*.

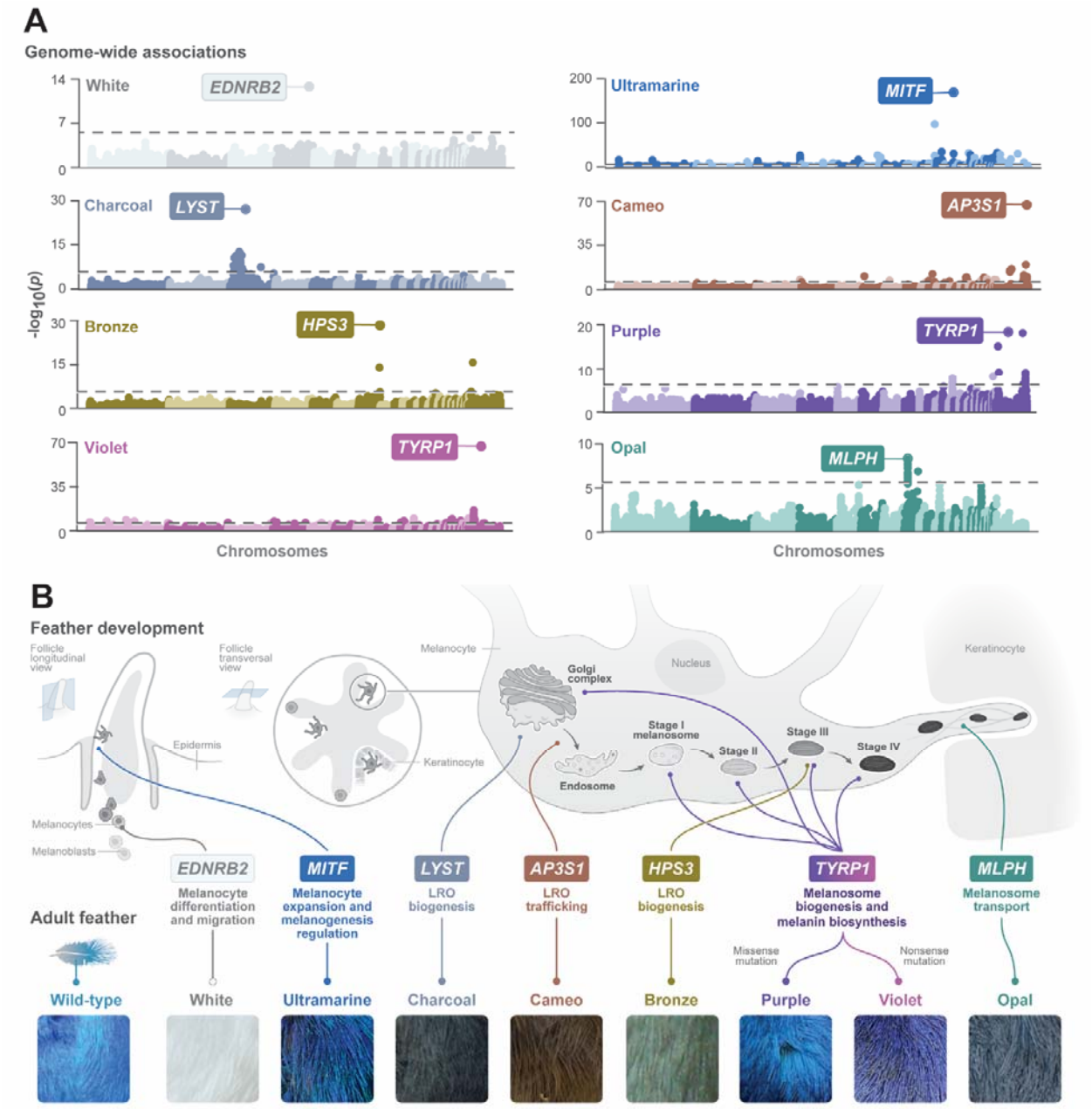

Figure 2. Genetic mapping of color variants. (A) Manhattan plots representing the genome-wide association analyses (GWAS) for plumage color. Each point represents a variant, with the *x*-axis showing genomic position (chromosomes in alternating colors, ordered according to the Indian peafowl reference genome) and the *y*-axis indicating statistical significance as $-\log_{10}(p)$. The horizontal dashed line indicates the Bonferroni-corrected genome-wide significance threshold. The most significantly associated variant in each plot is annotated with the name of the gene in which it resides. **(B)** Schematic overview of melanogenesis in feather development, highlighting candidate genes identified by GWAS. On the top left, longitudinal and cross-sectional views of a developing feather follicle depict the spatial arrangement of melanocytes during feather formation. In the center and top right, a melanocyte is illustrated with its key cellular components involved in melanogenesis, the melanosome maturation stages, and the dendritic projections connecting to keratinocytes, where melanosomes are finally deposited. Colored lines link genes to their respective subcellular locations and cellular processes associated with melanocytes, lysosome-related organelles (LRO), and melanosomes. *TYRP1* is associated with two phenotypes and the mutation type (e.g., missense, nonsense) is indicated. Along the bottom, magnified images of adult neck feathers (~2 × 2 cm) display the macroscopic color phenotypes corresponding to each mutation.

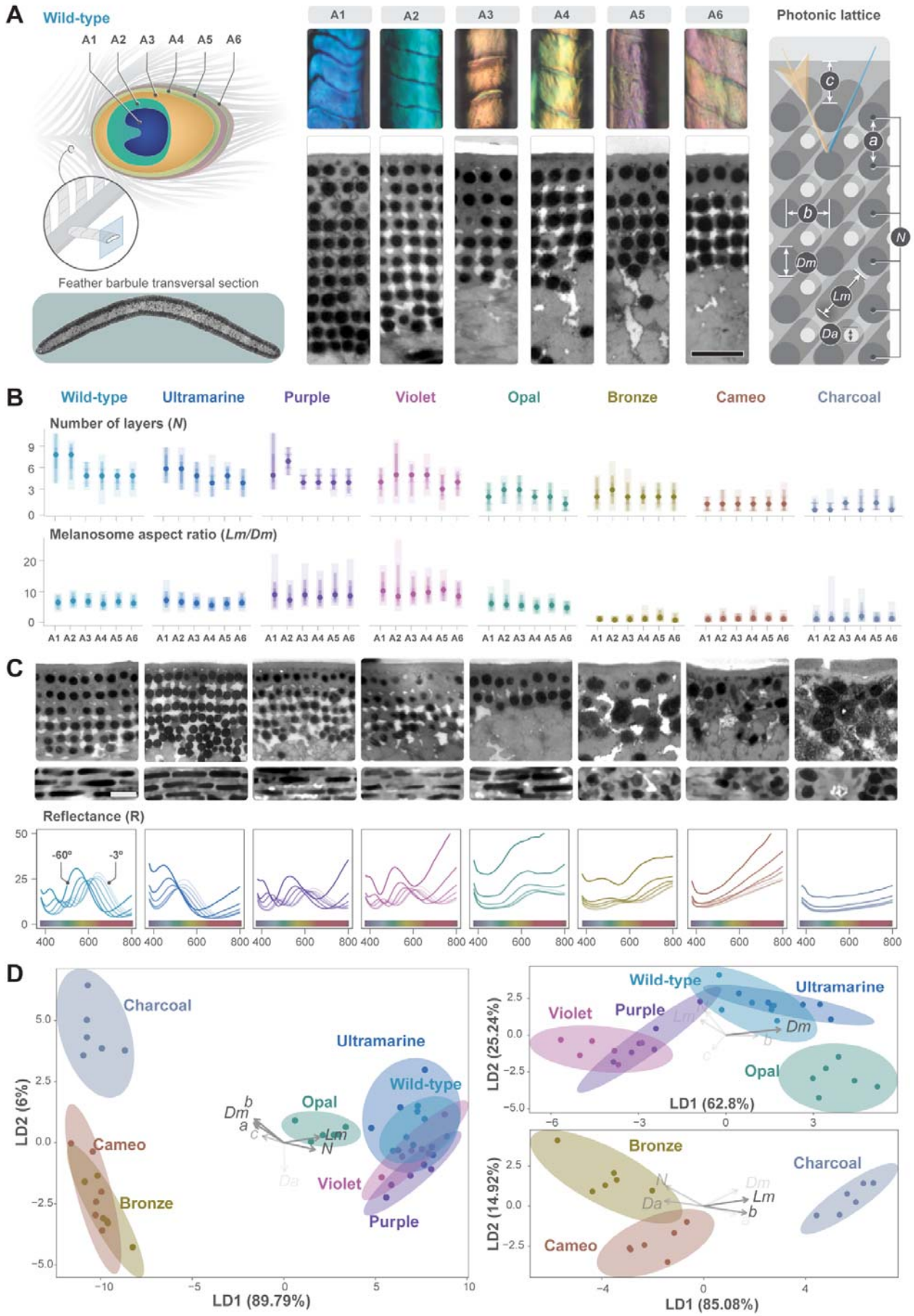

Figure 3. Ultrastructural modifications and photonic impact in eyespot feathers. (A) Schematic representation of feather structure, sampling procedure, and measured parameters of the photonic lattice.

Left: A diagram of the wild-type peacock eyespot, divided into six chromatic regions (A1–A6), alongside a transverse section of a feather barbule exemplifying the transmission electron microscopy (TEM) imaging used for structural analysis. Center: TEM images illustrating the ultrastructural arrangement of melanosomes across the six regions corresponding to different barbule colors in wild-type feathers (scale = 500 nm). Right: a 3D schematic representation of the key structural measurements from the photonic lattice of feather barbules, including: distance between melanosome layers (*a*); distance between melanosomes within layer (*b*); cortex thickness (*c*); melanosomes diameter (*Dm*); air pocket diameter (*Da*); number of melanosome layers (*N*) and length of melanosomes (*Lm*) (scale bar = 500 nm)**. (B)** Gradient plots illustrate differences in melanosome layer number (*N*) and aspect ratio (*Lm*/*Dm*) across A1–A6 in wild-type and mutants. Each circle represents an average measurement for a specific region, narrow bars represent quantile intervals at the 80% and 95% levels, and the wider bar represents the range of minimum and maximum for each area. Additional parameters are presented in **fig. S5**. **(C)** TEM of transversal (top) and longitudinal (bottom) cuts of feather barbules in A3 in wild-type and seven mutants (scale bar = 500 nm). Images for the remaining eyespot regions are available in **fig. S4**. Below, reflectance (R) profiles of A3 and variation (minimum–maximum) of the peak reflectance wavelength of the angle-dependent measurement from -60º to -3º (in increasing transparency). Measurements were taken at 10° intervals, with the exception of an additional measurement at −3° to capture the reflection near the surface normal. **(D)** Linear discriminant analyses (LDA) of structural measurements for all mutants on the left, and for iridescent (top) and non-iridescent (bottom) mutants on the right, with oval shading representing the variation per mutant across areas. The arrows represent the parameters explaining a significant amount of variance in LD1 of each analysis ($P < 0.01$), and the shading reflects their individual weight.

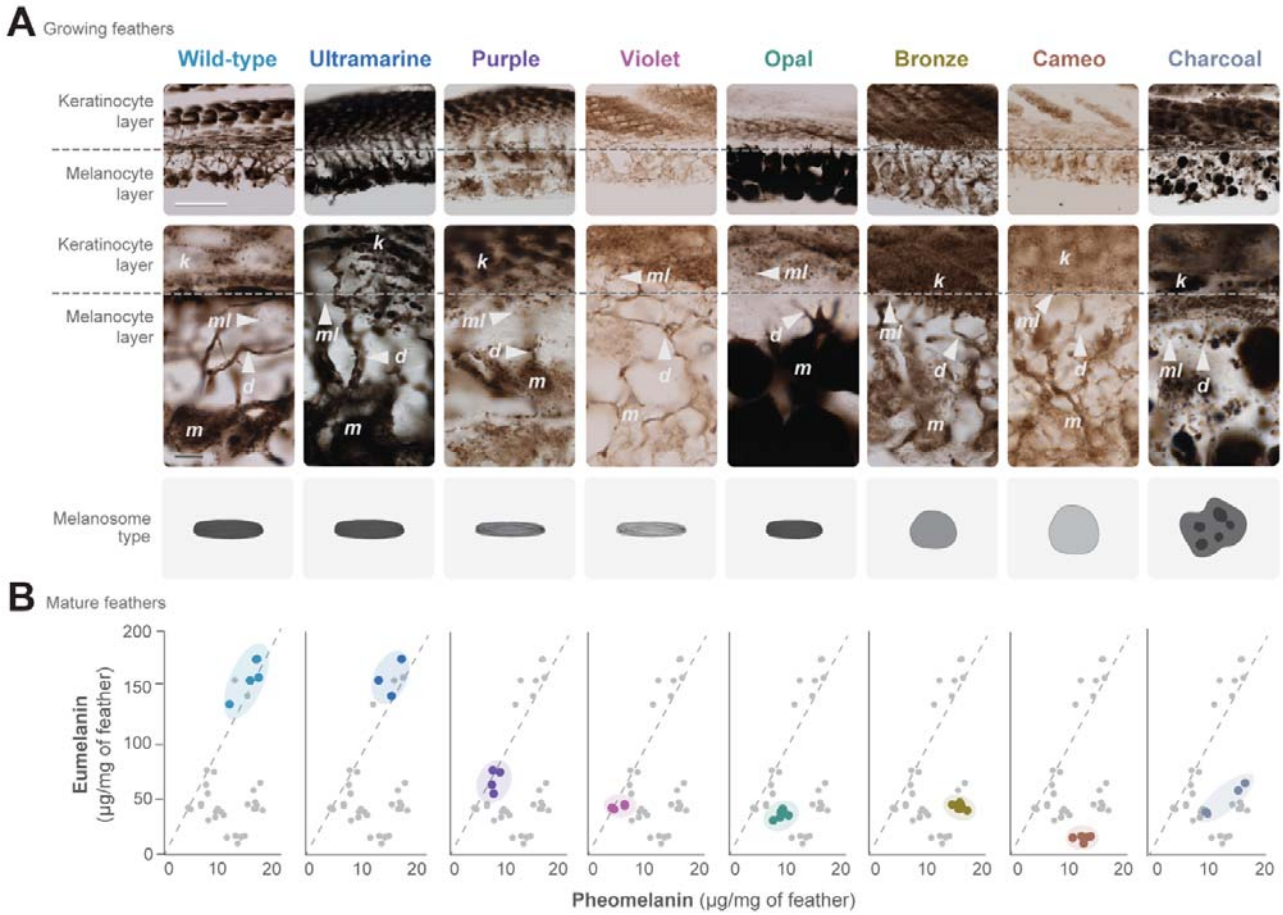

Figure 4. Histological and biochemical characterization of peafowl mutants. (**A**) Longitudinal sections of developing feather follicles of wild-type and mutant birds showing the organization of a melanocyte layer below the dashed line and a keratinocyte layer above the dashed line (top, scale bar = 50 µm). An enlarged view of the boundary between the two layers is shown (bottom, scale bar represents 10 µm), and various structures are highlighted: melanocytes (*m*), keratinocytes (*k*), dendrites (*d*), and melanosomes (*ml*). Across the bottom are schematic illustrations of melanosome morphology derived from TEM images (**fig. S4**). (**B**) Melanin quantification in mature neck feathers from wild-type and mutant phenotypes (detailed values provided in **table S6**). Graphs show quantification of pheomelanin (*x*-axis) and eumelanin (*y*-axis). All datapoints across all genotypes are plotted in each panel to enable direct visual comparison. Points belonging to the genotype shown in that panel are highlighted in color, while all other datapoints are shown in gray. The dashed line represents a eumelanin-to-pheomelanin ratio of 10:1.

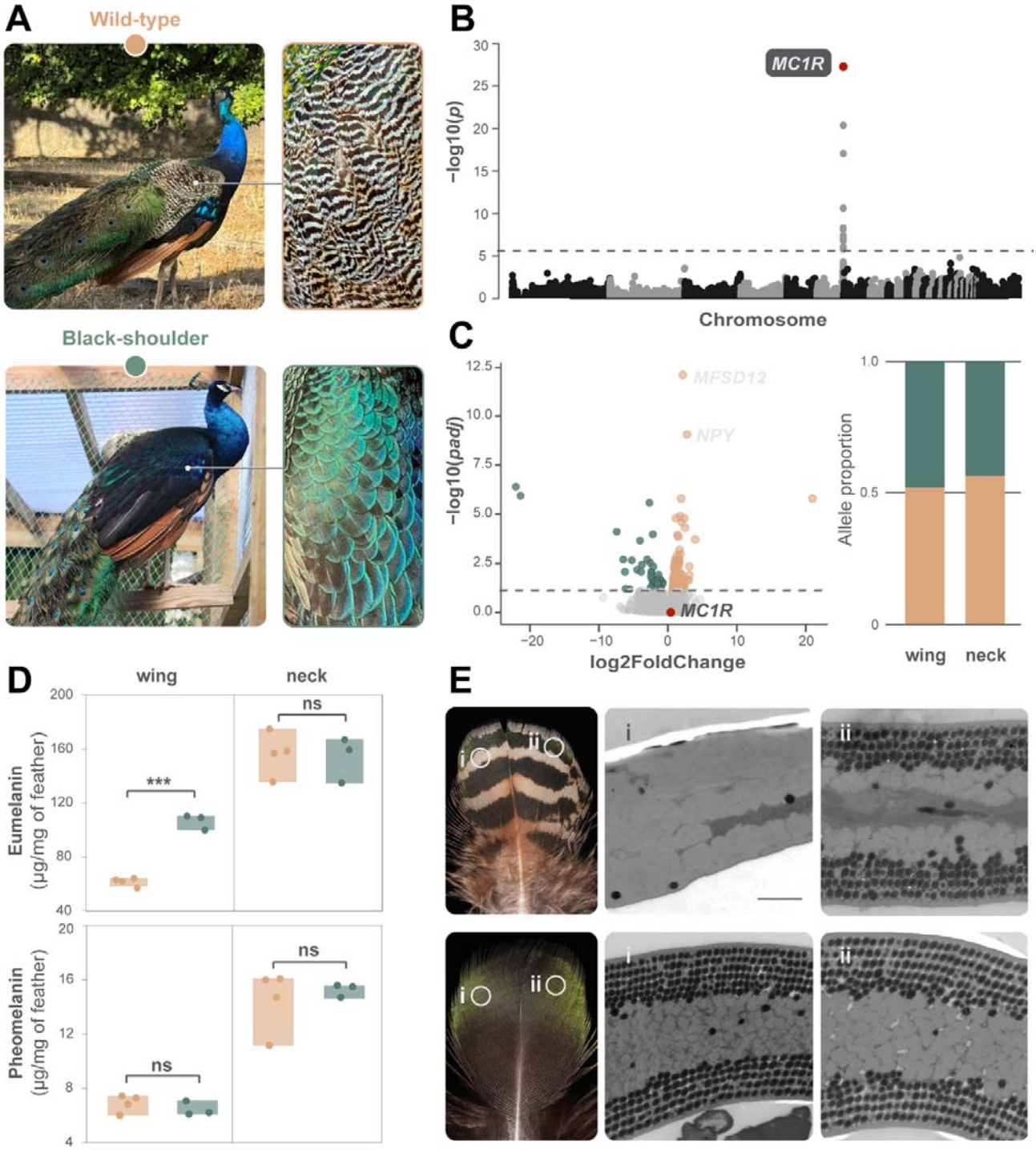

Figure 5. A patch-specific gain of iridescence. (A) Pictures illustrating wild-type (top) and *black-shoulder* (bottom) peacocks and a close-up of their wing covert feathers. **(B)** Manhattan plot displaying whole-genome association results. Each dot represents the $-\log_{10}(p)$ value from the likelihood ratio test for a genetic variant. The red horizontal line indicates the Bonferroni-corrected genome-wide significance threshold. The strongest association corresponds to a missense mutation in *MC1R* (red circle). **(C)** Left panel: a volcano plot comparing RNA-seq expression profiles of regenerating feather follicles from the wing of wild-type (n = 4) and *black-shoulder* (n = 4) individuals. The horizontal dashed line marks the significance threshold (FDR = 0.1). Green and beige circles represent genes upregulated in black-shoulder and wild-type individuals, respectively. The red circle represents the *MC1R* gene. Significantly differentially expressed genes with known functions in melanogenesis are labeled in grey. Right panel: allelic imbalance analysis of the wild-type (beige) and *black-shoulder* (green) *MC1R* alleles in feather follicles from the wing and neck of individuals heterozygous for the missense mutation associated with the *black-shoulder* phenotype. **(D)** Comparisons of eumelanin (top) and pheomelanin (bottom) levels in wing and neck feathers

of wild-type (beige) and *black-shoulder* (green) feathers (detailed values provided in **table S6**). Significant differences are indicated (*** $P < 0.001$), while "ns" denotes non-significant differences ($P > 0.05$), as determined by a two-sample *t*-test assuming unequal variances. **(E)** On the left, representative wing covert feathers from wild-type (top row) and *black-shoulder* (bottom row) individuals, highlighting equivalent locations chosen for ultrastructural analyses of feather barbules. The center and right panels show TEM images of transversal cuts of feather barbules from the selected regions, illustrating melanosome organization, and where scale represents 1 µm and is maintained across figures.

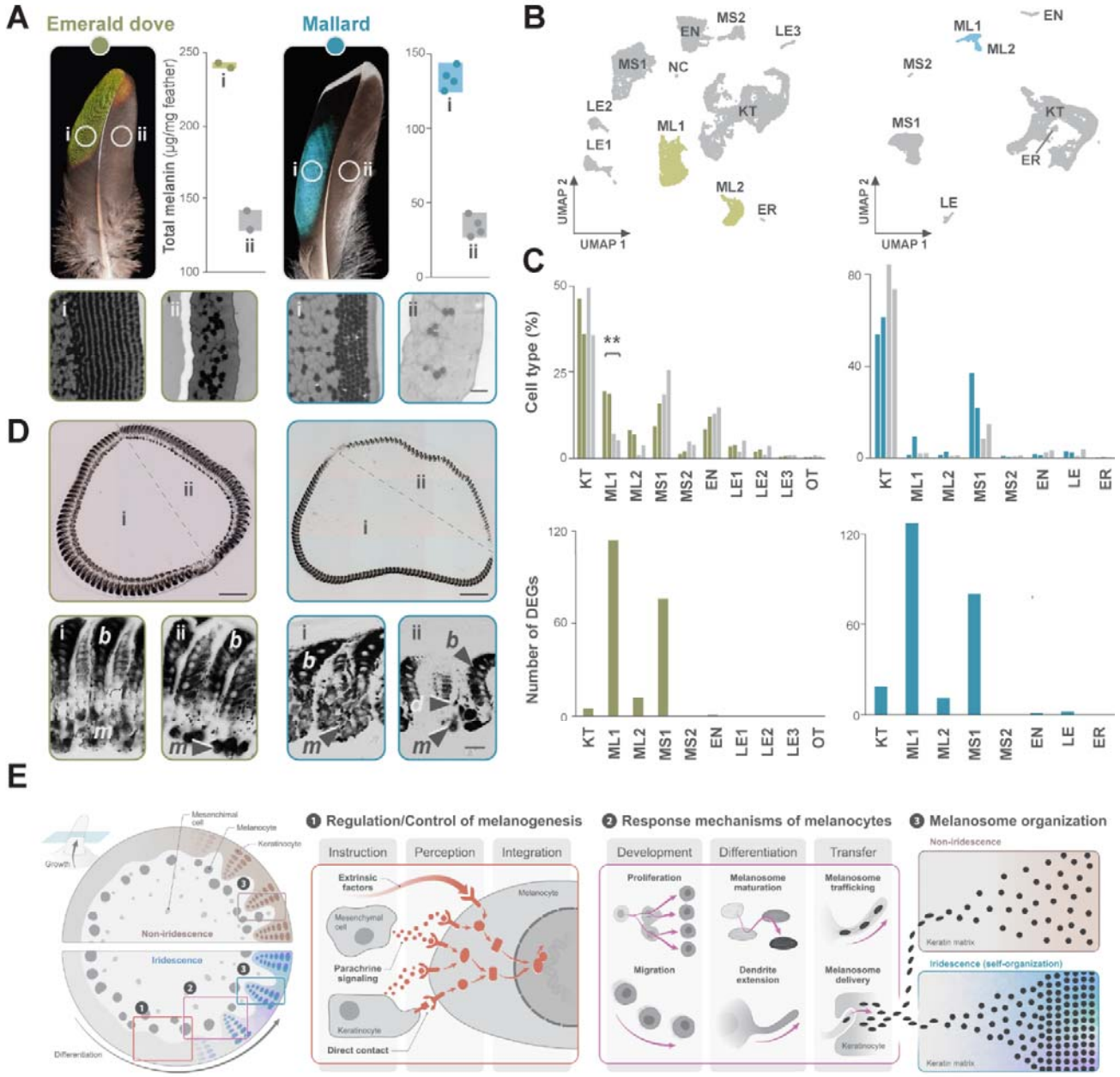

Figure 6. Modulation of iridescence via melanogenesis in wild bird species. The emerald dove (*Chalcophaps indica*) and the mallard duck (*Anas platyrhynchos*) develop asymmetric feathers containing both iridescent (i) and non-iridescent (ii) barbules. **(A)** *Top:* total melanin content per sample in each of the two feather regions. *Bottom:* TEM images of barbule cross-sections in iridescent and non-iridescent barbules (scale bar for all panels: 1 µm). **(B–C)** *scRNA-seq* of developing feathers from emerald dove (left) and mallard (right): ML, melanocytes; KT, keratinocytes; MS, mesenchymal cells; EN, endothelial cells; LE, leukocytes; ER, erythrocytes; NC, neural crest-derived cells; OT, others (ER+NC) **(B)** UMAP projections showing cell clustering. **(C)** *Top:* proportion of major cell types identified. *Bottom:* number of differentially expressed genes (DEGs) per cell type. **(D)** *Histological analyses* of developing feathers. *Top:* cross-sections of feathers showing overall morphology (scale bars: left = 200 µm; right = 500 µm). The dashed line indicates the boundary within the regenerating feather follicle separating regions that give rise to iridescent (i) and non-iridescent (ii) barbules. *Bottom:* high-magnification images of barb ridges (scale = 20 µm) reveal enhanced melanization associated with iridescence. m, melanocyte; d, dendrite; b, barbule keratinocyte. **(E)** Schematic summary of proposed mechanisms controlling iridescent feather color formation inferred from the scRNA-seq data. Cross-section representation of growing feathers developing non-iridescent (top) or iridescent (bottom) colors. Throughout barbule differentiation, depicted as a gradient from left to right, melanocytes receive, perceive, and integrate environmental cues from neighboring cells and extrinsic factors (e.g., paracrine signals or growth factors) (1). These instructions modulate melanocyte behavior (2), and the

resulting availability, shape, composition, and transfer of melanosomes influence their assembly into photonic structures (3).