RESEARCH ARTICLE SUMMARY

PIGMENTATION

A molecular mechanism for bright color variation in parrots

Roberto Arbore†*, Soraia Barbosa†, Jindřich Brejcha†, Yohey Ogawa, Yu Liu, Michaël P. J. Nicolaï, Paulo Pereira, Stephen J. Sabatino, Alison Cloutier, Emily Shui Kei Poon, Cristiana I. Marques, Pedro Andrade, Gerben Debruyn, Sandra Afonso, Rita Afonso, Shatadru Ghosh Roy, Uri Abdu, Ricardo J. Lopes, Peter Mojzeš, Petr Maršík, Simon Yung Wa Sin, Michael A. White, Pedro M. Araújo*, Joseph C. Corbo*, Miguel Carneiro*

INTRODUCTION: Coloration is an important trait in ecological adaptation and communication among animals, particularly birds, which use their diverse plumage for many purposes such as camouflage or social signaling. Parrots, known for their vivid coloration, display a wide range of hues including yellows, oranges, reds, and greens. This vibrant palette is primarily due to the deposition of psittacofulvins during feather growth, a class of pigments found exclusively in parrots.

RATIONALE: Psittacofulvins are polyene pigments that are endogenously synthesized to produce bright yellow, orange, and red colors. Combined with blue hues produced by feather nanostructures, yellow psittacofulvins are also essential for producing green colors. Although previous studies in domesticated species identified a polyketide synthase required for psittacofulvin biosynthesis, the mechanisms by which parrots diversify their color palette were previously unknown. The present study elucidates how psittacofulvins are biochemically modified to produce the broad spectrum of colors observed in wild parrot species and identifies the chemical and genetic bases of these color variations.

RESULTS: By combining spectroscopy, chromatography, and mass spectrometry analyses of feathers from various species, we uncovered a common chemical basis for yellow-to-red color variation in parrots. We found that the oxidation state of the psittacofulvin "end group" plays a key role in color shifts, with the tuning of color from yellow to red correlating with the ratio of carboxyl to aldehyde end group in psittacofulvin molecules; red feathers have large amounts of aldehyde psittacofulvins, whereas yellow and green feathers have higher levels of carboxyl psittacofulvins. To explore the genetic basis of these color differences, we studied the dusky lory, which occurs in two varieties in wild populations: yellow and red. Genetic mapping identified a genomic region associated with color variation, containing a candidate point mutation in a noncoding region downstream of the *ALDH3A2* gene, which encodes an enzyme that catalyzes the oxidation of fatty aldehydes to carboxylic acids. Single-cell RNA sequencing and chromatin accessibility assays in regenerating feather follicles confirmed that *ALDH3A2* is expressed at higher levels in late-differentiating keratinocytes—cells crucial for psittacofulvin metabolism—and reveal that the candidate causal mutation in the dusky lory lies within an open chromatin region specific to these keratinocytes, suggesting that the causal variant affects the activity of an enhancer which controls levels of *ALDH3A2* expression in a cell type–specific manner. Transcriptomic analyses in dusky lory and another parrot species (rosy-faced lovebirds) indicate that regenerating feather patches enriched for yellow psittacofulvins have higher levels of *ALDH3A2* than patches enriched for red psittacofulvins. Yeast assays confirmed that the ALDH3A2 enzyme is capable of converting red psittacofulvins into yellow pigments as predicted by the genetic results.

CONCLUSION: This study identifies ALDH3A2 as a key enzyme in the biochemical pathway responsible for color variation in parrots. These results reveal insights into the molecular mechanisms underlying one of the most visually striking adaptations in the natural world and lay the groundwork for future studies aimed at understanding how bright colors evolve in the wild. ■

The full list of author affiliations can be found in the full article online.
*Corresponding author. Email: miguel.carneiro@cibio.up.pt (M.C.); jcorbo@wustl.edu (J.C.C.); pmaraujo.ecotop@gmail.com (P.M.A.); roberto.arbore@cibio.up.pt (R.Ar.)
†These authors contributed equally to this work.
Cite this article as R. Arbore *et al.*, *Science* **386**, eadp7710 (2024). DOI: 10.1126/science.adp7710

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adp7710

The molecular bases of bright color variation in parrots. Yellow-to-red color variation in parrot feathers is due to differences in the concentration of yellow carboxyl and red aldehyde psittacofulvin pigments. Through a combination of genetic and biochemical techniques, we identify aldehyde dehydrogenase 3 family member A2 (*ALDH3A2*) as a key enzyme regulating the balance of aldehyde to carboxyl pigments in parrots.

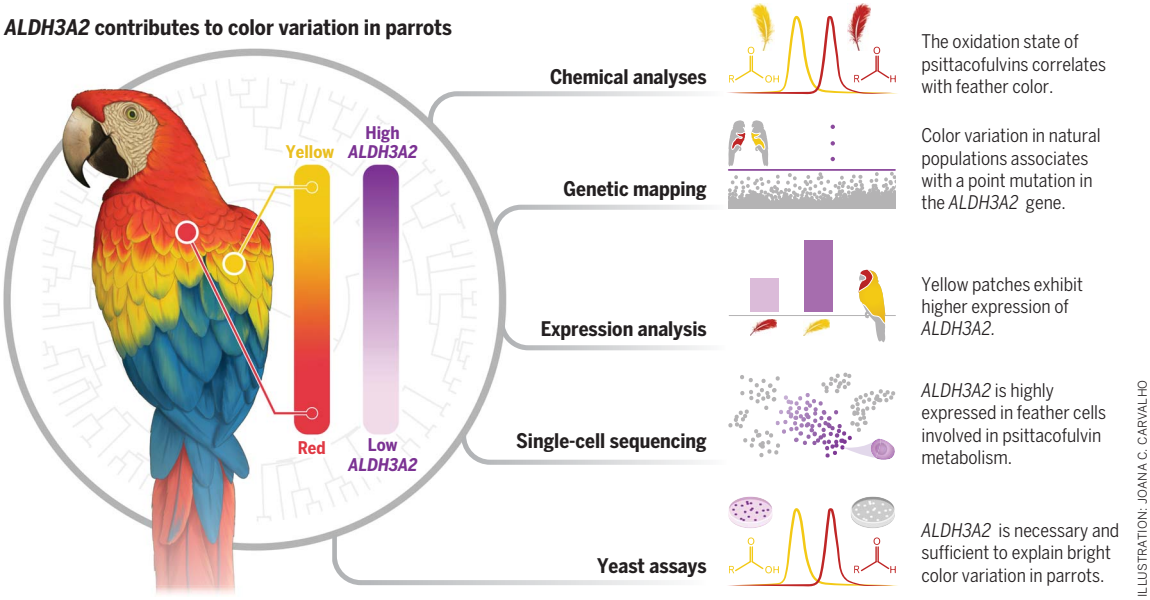

ILLUSTRATION: JOANA C. CARVALHO

RESEARCH ARTICLE

PIGMENTATION

A molecular mechanism for bright color variation in parrots

Roberto Arbore[1,2,3]†*, Soraia Barbosa[1,2]†, Jindřich Brejcha[4]†, Yohey Ogawa[3], Yu Liu[3], Michaël P. J. Nicolaï[5,6], Paulo Pereira[1,2,7], Stephen J. Sabatino[1,2], Alison Cloutier[8], Emily Shui Kei Poon[8], Cristiana I. Marques[1,2,7], Pedro Andrade[1,2], Gerben Debruyn[5], Sandra Afonso[1,2], Rita Afonso[1,2,7], Shatadru Ghosh Roy[9], Uri Abdu[9], Ricardo J. Lopes[1,2,10,11], Peter Mojzeš[12], Petr Maršík[13], Simon Yung Wa Sin[7], Michael A. White[14,15], Pedro M. Araújo[1,2,16]*, Joseph C. Corbo[3]*, Miguel Carneiro[1,2]*

Parrots produce stunning plumage colors through unique pigments called psittacofulvins. However, the mechanism underlying their ability to generate a spectrum of vibrant yellows, reds, and greens remains enigmatic. We uncover a unifying chemical basis for a wide range of parrot plumage colors, which result from the selective deposition of red aldehyde- and yellow carboxyl-containing psittacofulvin molecules in developing feathers. Through genetic mapping, biochemical assays, and single-cell genomics, we identified a critical player in this process, the aldehyde dehydrogenase *ALDH3A2*, which oxidizes aldehyde psittacofulvins into carboxyl forms in late-differentiating keratinocytes during feather development. The simplicity of the underlying molecular mechanism, in which a single enzyme influences the balance of red and yellow pigments, offers an explanation for the exceptional evolutionary lability of parrot coloration.

Colors play a vital role in ecological adaptation and communication in the natural world (*1*). Among animals, birds stand out with their wide range of striking hues, color patterns, and iridescence. Through their plumage colors, birds interact with their environment and convey crucial information about individual and species identity, health status, sexual attractiveness, and social dominance (*1*–*6*). Bird colors are thus frequent targets of both natural and sexual selection (*7*, *8*). Despite decades of study, understanding the selective and ecological pressures underlying the adaptive function of coloration in nature (*9*), as well as the physiological and metabolic processes presumably linking color ornaments to fitness, remains a challenge (*1*, *10*). Investigating the molecular mechanisms controlling color variation is one promising approach to answer some of these enduring questions.

Parrots are renowned for their vibrant plumage (*11*, *12*). Their feathers differ considerably among species in hue, saturation, and overall patterning across the body (Fig. 1) and likely serve a variety of signaling and nonsignaling functions (*13*–*18*). The rapid and dynamic evolution of parrot coloration (Fig. 1) largely results from the differential deposition of psittacofulvins during feather growth (*12*, *19*–*21*), a class of polyene pigments that are only found in these birds and which generate their fiery reds and luminous yellows (*19*, *22*). When combined with blue coloration that arises from light scattering by nanostructural features of the feather, yellow psittacofulvin coloration also gives rise to vivid green hues (Fig. 1). Unlike carotenoids, which also produce bright yellow and red colors in many other bird species and need to be acquired through diet (*23*–*25*), psittacofulvins are endogenously synthesized by parrots (*22*). A polyketide synthase (PKS) has been identified as essential for psittacofulvin biosynthesis in domesticated mutants lacking psittacofulvin-based pigmentation (*26*–*28*), but the mechanisms governing the color variation found in parrots remain unknown.

In this study, we identify a simple molecular mechanism that explains how psittacofulvins are biochemically modified to produce yellow-to-red and green hues in parrots. The discovery of this mechanism provides an explanation for a broad spectrum of phenotypic variation that characterizes one of the most brilliantly colored animal groups in the natural world.

A shared chemical basis for yellow-to-red psittacofulvin coloration in parrots

Psittacofulvins were first identified in studies of red parrot feathers as noncarotenoid pigment molecules consisting of an extended polyene chain with a single aldehyde end group (Fig. 2A) (*19*, *22*). Subsequent research postulated that variation in the hue of psittacofulvin-based colors could arise from multiple factors (*22*, *26*, *29*); however, the precise chemical and physical mechanisms responsible for color variation in parrots remain unclear. To explore these questions, we conducted comprehensive chemical analyses of red, orange, yellow, and green feathers across representatives of all parrot superfamilies (seven species), spanning >50 to 80 million years of evolution (*30*).

We utilized confocal Raman microscopy to examine differences in the vibrational spectra of pigment molecules in situ. These analyses revealed a consistent tendency of yellow/green hues to be shifted toward higher wavenumbers compared with red hues, regardless of the species analyzed (Fig. 2B and figs. S1 and S2), aligning with previous studies (*21*, *29*, *31*). We also confirmed that green color patches result from a combination of blue structural coloration and yellow psittacofulvin pigments and therefore exhibit Raman spectra similar to those produced by yellow patches. These analyses further revealed that yellow/green and red hues share a common structural fingerprint across species and point toward a general mechanism underlying psittacofulvin-based color differences among parrots.

To gain further insights into pigment composition in parrot feathers, we performed ultra-high-performance liquid chromatography (UHPLC) coupled with ultraviolet/visible (UV/VIS) detection and high-resolution, accurate-mass (HRAM) quadrupole time-of-flight (Q-TOF) mass spectrometry. Our chemical analyses confirmed the existence of psittacofulvins with varying polyene chain lengths (14, 16, or 18 carbons; herein referred to as C14, C16, and C18, respectively), and two different types of functional groups: aldehydes and carboxylic acids (Fig. 2C and figs. S3 and S4). These various molecular forms were identified in all species and feather tracts. We found that both the number of conjugated double bonds and the identity of the end group

[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]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA. [4]Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czech Republic. [5]Evolution and Optics of Nanostructures Group, Biology Department, Ghent University, Ghent, Belgium. [6]Department of Recent Vertebrates, Royal Belgian Institute of Natural Sciences, Brussels, Belgium. [7]Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal. [8]School of Biological Sciences, The University of Hong Kong, Hong Kong. [9]Department of Life Sciences, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel. [10]MHNC-UP, Natural History and Science Museum of the University of Porto, Porto, Portugal. [11]cE3c – Center for Ecology, Evolution and Environmental Change & CHANGE, Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal. [12]Institute of Physics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. [13]Department of Food Science, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Prague, Czech Republic. [14]Edison Family Center for Systems Biology and Genome Sciences, Washington University School of Medicine, St. Louis, MO, USA. [15]Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA. [16]University of Coimbra, MARE – Marine and Environmental Sciences Centre, Department of Life Sciences, Coimbra, Portugal.
***Corresponding author. Email: miguel.carneiro@cibio.up.pt (M.C.); jcorbo@wustl.edu (J.C.C.); pmaraujo.ecotop@gmail.com (P.M.A.); roberto.arbore@cibio.up.pt (R.Ar.)**
†These authors contributed equally to this work.

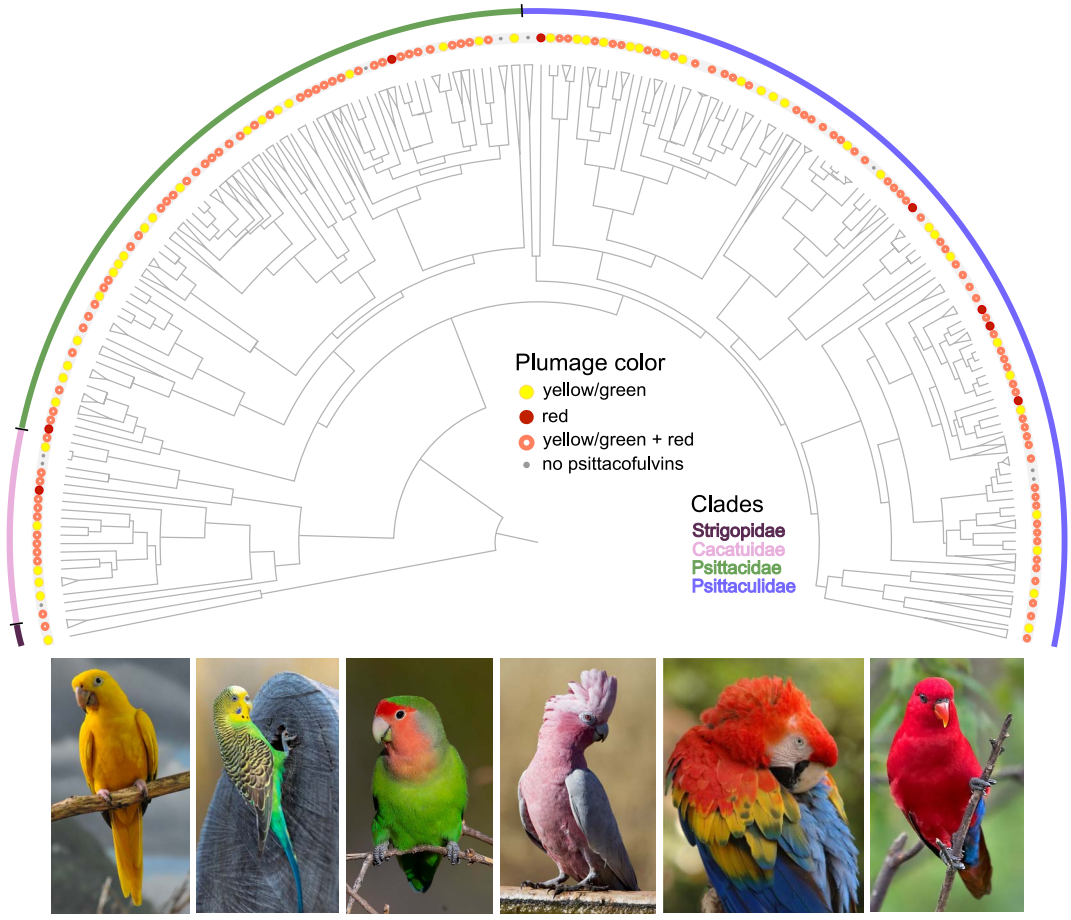

Fig. 1. Psittacofulvin-based coloration diversity and evolution in parrots. Phylogeny of all 354 species of parrots showing whether species display (i) only yellow/green hues (yellow circles); (ii) only red hues (red circles); (iii) simultaneously yellow/green and red hues (orange circles); or (iv) no psittacofulvin-based hues (grey circles, i.e., color is produced by other mechanisms with no contribution from psittacofulvins). We note that green and yellow color patches were considered together since the green color is a combination of blue structural coloration and yellow psittacofulvin pigments. Lineages including several species of identical phenotype are collapsed. The phylogeny demonstrates the high number of evolutionary shifts among parrots expressing yellow, green, and red. Across the bottom is a compilation of photographs showcasing the diversity of parrot plumage coloration. (From left to right) golden parakeet (*Guaruba guarouba*, CC BY-SA 3.0 Rodrigo Menezes), budgerigar (*Melopsittacus undulatus*, ML619272371, Robert Hynson), rosy-faced lovebird (*Agapornis roseicollis*, ML215809581, Niall Perrins), galah (*Eolophus roseicapilla*, CC BY-SA 2.0 Jim Bendon), scarlet macaw (*Ara macao*, ©Milan Kořínek, https://biolib.cz/en), and red lory (*Eos bornea*, CC BY-SA 3.0 Doug Janson).

influence the wavelength of maximal absorbance of the psittacofulvin molecule (Fig. 2C). As with carotenoid pigments (*24*), the addition of a single double bond to the conjugated system of a psittacofulvin redshifts its peak absorbance by 16 to 20 nm. Similarly, as reported for retinoids (*32*), switching from a carboxylic acid to an aldehyde end group redshifts the peak absorbance by 17 to 24 nm.

In our analyses, red and orange feathers were strongly enriched in aldehyde psittacofulvins whereas yellow and green feathers contained a higher proportion of carboxyl psittacofulvins (F = 43.8, P = 8.94 × 10^{-16}; Fig. 2D and table S1). This pattern was consistently observed across all major parrot lineages (Fig. 2D and fig. S5). Intensely red-colored patches also exhibited a tendency for a higher proportion of molecules with longer carbon chain lengths, as well as a higher concentration of pigments (e.g., macaw, kea, and Pesquet's parrot; Fig. 2D). However, the same patterns were not evident in less-saturated red or orange patches (e.g., cockatiel and lovebird). Taken together, these findings suggest that the ratio of carboxylic acid- to aldehyde-containing psittacofulvins plays a major role in determining the hue of a feather and that this mechanism is conserved across divergent lineages of parrots.

The genetic and transcriptomic bases of psittacofulvin coloration

Heritable differences in psittacofulvin coloration are largely fixed among parrot species. The absence of phenotypic variation within populations complicates efforts to study the genetic basis of parrot coloration, which has been exclusively analyzed in domesticated species carrying mutations that result in a complete loss of psittacofulvin pigmentation (*26*, *28*). To uncover molecular mechanisms underlying the evolution and diversification of psittacofulvin-based colors in nature, we took advantage of a naturally occurring intraspecific polymorphism present in the dusky lory (*Pseudeos fuscata*). This species is native to New Guinea, where two color morphs expressing red and yellow pigmentation coexist and interbreed in sympatry (Fig. 3A) (*33*), offering a rare opportunity to conduct genetic mapping and investigate how parrot colors evolve in the wild. The red and yellow morphs differ in the content of psittacofulvin forms as described above for the other parrot species (Fig. 3B and figs. S5 and S6). Pedigree and phenotypic data gathered from 20 breeding pairs show that this polymorphism is inherited largely as a binary trait, indicative of a simple genetic architecture, with yellow being dominant over red (table S2).

We assembled and annotated a draft reference genome of the dusky lory (tables S3 to S5) and resequenced the genomes of 57 individuals representing both color morphs (red n = 35; yellow n = 22; mean depth = 11.5 ± 4.8×; table S6). Among the 4,303,897 variants we examined through association tests, only three exceeded the genome-wide significance threshold for explaining the color phenotype (Fig. 3C). The three variants spanned a small interval of 284 base pairs (bp) (scaffold_13:6,288,645-6,288,929), and one single nucleotide polymorphism (SNP) (scaffold_13:6,288,712T>C) showed a markedly stronger association with color (P = 6 × 10^{-14}). Both alleles of the top associated variant were present in multiple haplotype backgrounds and exhibited a rapid decay of haplotype homozygosity [extended

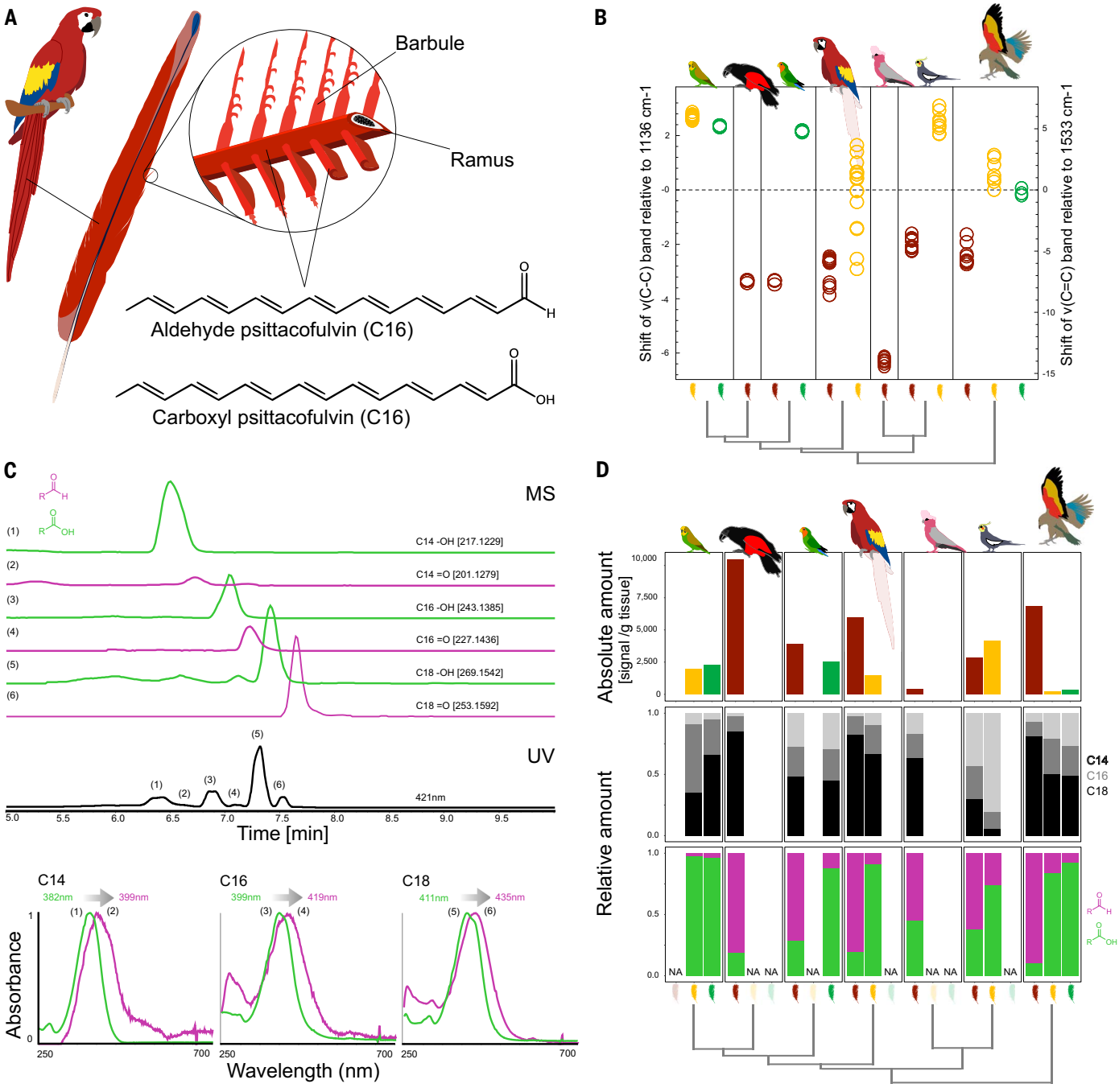

Fig. 2. Chemical analyses of psittacofulvin pigmentation. (**A**) Schematic representation of a tail feather of a scarlet macaw (*Ara macao*). Psittacofulvins are deposited within the keratin matrix of both the ramus and barbules of feathers (insert). Psittacofulvins are linear polyenes with various carbon chain lengths (C16 examples shown) and with distinct terminal groups (aldehyde or carboxyl groups). (**B**) The plot illustrates the magnitude of the shifts in the positions of the two primary Raman bands characteristic of psittacofulvins (both *y*-axis). The dashed line at 0 represents the average spectrum. The variation in these Raman bands is shown for yellow, green, and red feathers of the studied parrot species. (From left to right) budgerigar (*Melopsittacus undulatus*), Pesquet's parrot (*Psittrichas fulgidus*), rosy-faced lovebird (*Agapornis roseicollis*), scarlet macaw (*Ara macao*), galah (*Eolophus roseicapilla*), cockatiel (*Nymphicus hollandicus*), and kea (*Nestor notabilis*). (**C**) Chromatograms of the presence of positive ions of the exact molecular masses corresponding to psittacofulvins in the carboxyl (green) and aldehyde forms (magenta) detected by mass spectrometry (HRAM-QTOF). Peak 1 corresponds to C14 carboxylic acid, 2 to C14 aldehyde, 3 to C16 carboxylic acid, 4 to C16 aldehyde, 5 to C18 carboxylic acids, and 6 to C18 aldehyde. The mass spectrometry peaks correspond to the UV/VIS-detected absorbance peaks (UHPLC UV/VIS) shown below—the slight shift is caused by the delay from the UV/VIS detection to the HRAM-QTOF molecular mass detection. The UHPLC UV/VIS chromatogram (black) shows absorbance peaks detected at 421 nm, which is close to the average maximum absorbance wavelength of psittacofulvins. The plots at the bottom show the maximum absorbance shifts between the carboxyl and aldehyde forms of psittacofulvins with different carbon chain lengths. The absorbance spectra have been normalized such that the maximum intensity = 1. (**D**) Differences in psittacofulvin content in red, yellow, and green feathers of parrot species. The total amount of psittacofulvins

or the relative amount of each type was quantified as the area under the peak in the exact mass spectrum relative to the baseline. The upper row shows the total amount of psittacofulvins, the middle row shows the relative abundance of psittacofulvins of different lengths, and the lower row shows the ratio of aldehyde (magenta) and carboxylic forms (green). The order of the species is the same as in (B).

haplotype homozygosity (EHH) < 0.5, ~1.2 kb, and ~2 kb on each side; fig. S7]. These patterns imply that color alleles at this locus have been coexisting and recombining within dusky lory wild populations for many generations. Based on patterns of nucleotide variation around the candidate region, we inferred the color polymorphism to be ~1,000,000 years old (table S7), ranging from ~460,000 to ~1,940,000 years, depending on assumptions about generation time, mutation rate, and recombination rate (*34*).

The candidate interval lies in a non-coding region between *ALDH3A2* and *SLC47A1*, immediately downstream of the last exon of *ALDH3A2* (Fig. 3C). *SLC47A1* encodes the Solute Carrier Family 47 Member 1, a transporter involved in excreting endogenous and exogenous electrolytes through urine and bile (*35*). *ALDH3A2* encodes the Aldehyde Dehydrogenase 3 Family Member A2 (also known as FALDH), a ubiquitously expressed enzyme responsible for catalyzing the oxidation of medium- and long-chain fatty aldehydes (preferentially acting on C14 to C18 substrates) to the corresponding carboxylic acids (*36*). In light of our pigment analyses, *ALDH3A2* is a strong candidate gene for explaining the differences in psittacofulvin coloration between dusky lory color morphs.

To investigate whether color differences between the dusky lory color morphs could arise from differences in gene expression, we performed bulk RNA-seq analysis and generated full-length transcriptomes through Pacific Biosciences (PacBio) long-read sequencing of regenerating feather follicles derived from both red (n = 3) and yellow (n = 3) individuals (table S8). Our decision to examine growing feathers was guided by the discovery that parrots neither circulate psittacofulvins in the bloodstream nor accumulate them in the liver (*22*), implying that pigment biosynthesis occurs locally in the integument during feather development. We identified 33 genes with significant differential expression, but none were contained within the candidate scaffold (data S1). Considering the positional information provided by our genetic mapping, we examined the two genes flanking the associated variants. *SLC47A1* displayed negligible expression in regenerating feather follicles. For *ALDH3A2*, we found that the three protein-coding isoforms detected by Iso-seq (fig. S8) are expressed at similar levels in red and yellow birds. Although we detected a subtle, but statistically nonsignificant, trend toward higher *ALDH3A2* expression in yellow individuals by RNA-seq (Fig. 3D; left), quantitative polymerase chain reaction (qPCR) analyses failed to confirm this difference (fig. S9).

Differences in the expression of *ALDH3A2* between color morphs may be obscured by its ubiquitous expression (see below), insufficient statistical power due to low sample size, or by the fact that all three yellow individuals under investigation were heterozygous for the *red* and *yellow* alleles at the candidate locus. To further explore *ALDH3A2* expression, we compared the relative abundance of *red* and *yellow* transcripts in each of the heterozygous individuals using Iso-seq data. This approach is particularly sensitive to minor differences in expression given that in heterozygous birds both alleles are exposed to the same trans-acting regulatory environment in the nucleus. We found an imbalance in the expression of the two alleles, with a higher percentage of reads corresponding to the *yellow* allele (71% *yellow* vs. 29% *red*, χ^2, P = 0.03; Fig. 3D; right), suggesting the existence of underlying cis-regulatory differences favoring higher expression of this allele. This finding is consistent with the hypothesis that *ALDH3A2* encodes an enzyme that converts aldehyde psittacofulvins into carboxyl forms and fits the expectations of our pigment analysis (Fig. 3B), which showed that dusky lory individuals with yellow plumage contained a higher proportion of carboxyl psittacofulvins in their feathers.

The above hypothesis implies that differences in the ratio of carboxyl- to aldehyde-containing psittacofulvins across feathers should positively correlate with *ALDH3A2* expression levels. To test this idea, we used rosy-faced lovebirds (*Agapornis roseicollis*), a species that displays both green (i.e., yellow psittacofulvin–containing) and red plumage patches (Fig. 1). Using bulk RNA-seq, we analyzed gene expression in regenerating feather follicles of eight rosy-faced lovebirds and four plumage patches: three green patches (from back, chest, and head) and one red feather patch (from head) (table S8). Our initial chemical analyses indicate that these patches differ in the ratio of carboxyl- to aldehyde-containing psittacofulvins, with green patches showing a higher ratio than red (Fig. 2D). We found that *ALDH3A2* ranks among the top differentially expressed genes in all pairwise comparisons between green and red patches (fig. S10), with 4 to 6 times higher expression in the three green feather patches compared with the red (Fig. 3E). Differential expression of *ALDH3A2* between color patches was confirmed by qPCR (fig. S9). These findings are consistent with the hypothesis that yellow/green feathers express higher levels of *ALDH3A2*, which converts red aldehyde psittacofulvins into the corresponding yellow carboxyl forms. These data further support the conclusion that *ALDH3A2* is a strong candidate gene for mediating color variation in parrots.

ALDH3A2 is expressed at higher levels in late-differentiating keratinocytes

To investigate the expression of *ALDH3A2* in the context of feather development, we next studied gene expression at the cellular level. We generated single-cell RNA sequencing (scRNA-seq, n = 2) data from regenerating feather follicles of budgerigars (table S9), a parrot species that expresses yellow psittacofulvin–based coloration (Figs. 1 and 2) and can be maintained for extended periods under laboratory conditions required for optimization of single-cell analysis protocols. We analyzed a total of 6262 cells, which aggregated in 10 clusters representing the major cell types of regenerating feather follicles (Fig. 4A and fig. S11) (*34*). Distally located epithelial and pulp cells undergo apoptosis and keratinization during feather regeneration (*37*), hence we expect these cell populations to be underrepresented and for more immature, proximally placed cells to be correspondingly overrepresented in our single-cell dataset. We found that *ALDH3A2* exhibited widespread expression across all cell types (fig. S12), as expected for a gene critical for cellular metabolism (*38*).

The deposition of psittacofulvin pigments within the keratin matrix of feathers suggests that keratinocytes might be involved in their metabolism. In a subset of the cell clusters identified using scRNA-seq, we noticed a strong enrichment of various keratin genes such as *keratin 17-like* (*KRT17L*), a type-I alpha keratin known to have ubiquitous expression in feather keratinocytes (*39*) (Fig. 4B). In regenerating feather follicles, keratinocyte differentiation proceeds along the proximal-distal axis of growth from the proliferative zone to the distal tip of the feather and radially during the formation of barb ridges in which they organize into the marginal, axial, and barbule plates (*40*). A closer examination of cell cycle marker genes in our scRNA-seq data revealed a population of dividing keratinocytes with elevated expression of these genes, likely corresponding to the proliferative zone at the base of the feather follicle (Fig. 4C and fig. S12) (*34*). Additionally, gene expression patterns suggested a progression of keratinocyte lineages from actively dividing cells to noncycling late-differentiating keratinocytes, likely positioned toward the distal tip of the feather and enriched in expression of *SCEL* (*41*) (a precursor to the cornified envelope of

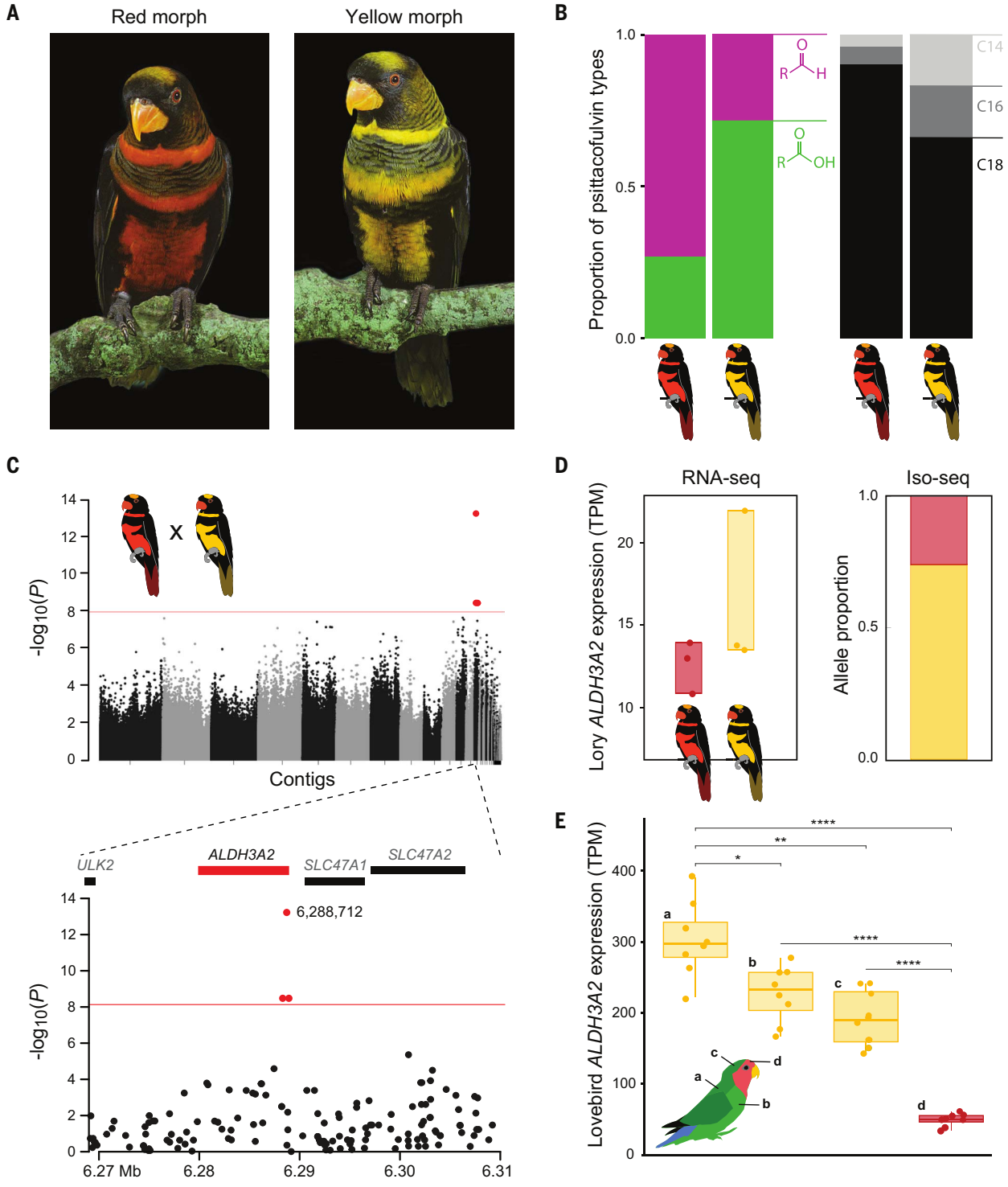

Fig. 3. The genetic and transcriptomic bases of psittacofulvin colors. (**A**) Images of red and yellow morphs of the dusky lory (*Pseudeos fuscata*). Photo credits: David Hosking/Minden Pictures. (**B**) Differences in pigment composition between feathers of red and yellow morphs. (Left) relative ratio of aldehydes (magenta) and carboxylic acids (green). (Right) relative abundance of psittacofulvins of different chain lengths. (**C**) Genetic mapping of the color polymorphism. The top panel summarizes the genome-wide association analysis using whole-genome resequencing data. Each dot represents the $-\log_{10}$ transformation of Wald test *P*-values for each variant. The horizontal red line indicates the Bonferroni-corrected genome-wide significance [$P = 1.16 \times 10^{-8}$; $-\log_{10}(P) = 7.94$] based on the total number of tests (n = 4,303,897). The bottom panel is a zoomed-in view of the region of association shown on top. The protein coding genes contained within the represented genomic interval are shown at the top of the panel. (**D**) Patterns of gene expression of *ALDH3A2* between dusky lory morphs. (Left) RNA-seq normalized raw read counts (circles) from regenerating feather follicles from red (n = 3, left) and yellow (n = 3, right) birds, with colored boxes illustrating the range of read counts for the respective color morph. (Right) proportion of full-length Iso-seq transcripts (n = 152) linked to the red and yellow alleles in heterozygous individuals (n = 3). (**E**) Differential expression of *ALDH3A2* between

red feathers from the forehead region (d) versus green feathers from the back (a), chest (b), and head (c) regions of rosy-faced lovebirds (n = 8 for each region; Welch's test followed by Games-Howell post-hoc test; * adj-P < 0.05; ** adj-P < 0.01; *** adj-P < 0.001; **** adj-P < 0.0001). The sampled regions are indicated on the illustration at the bottom left of the graph.

terminally differentiated keratinocytes) and of the epidermal differentiation complex gene *EDQM3* (*42*) (Fig. 4C and fig. S12). By explicitly modeling the progress of keratinocyte differentiation along a branching trajectory rooted at the putative follicle proliferative zone (*34*), we found a progressive increase in *ALDH3A2* expression toward late differentiating cells, with maximum expression in a population of cells expressing *NCAM1* (*43*) and likely representing axial plate keratinocytes (Fig. 4D and fig. S12). Axial plate keratinocytes are located in the midline of each developing barb ridge and will later undergo apoptosis, enabling the flanking barbule plate cells to open and form a feather barb (*37*). The increased expression of *ALDH3A2* in this lineage of late-differentiating keratinocytes suggests that they serve as the primary site for yellow-to-red psittacofulvin conversion in parrots during feather development.

A candidate causal mutation resides in a late-differentiating keratinocyte–specific open chromatin region

Next, we sought to identify the specific mutation responsible for regulating *ALDH3A2* expression in the dusky lory. The yellow morph is genetically dominant over the red morph, establishing a clear expectation for the genotypes associated with causative mutations. Apart from the three significant non-coding variants from the genetic mapping analyses (Fig. 3C), we did not identify any structural variants within 100 kb of the candidate interval that followed the expected inheritance pattern, nor did we identify additional small deletions or point mutations. The lead variant identified in the association mapping analysis (scaffold_13:6,288,712T>C), located in a non-coding region 42 bp downstream of the longest *ALDH3A2* transcript, exhibited the anticipated genotypes in nearly all individuals (98%). One yellow individual deviated from the expected pattern. However, the two remaining significant variants (scaffold_13:6,288,645C>T and 6,288,929A>G) did not match the expected genotypes based on phenotype in five individuals, including the same yellow individual that carried a *red* haplotype in homozygosity across the entire region. Additionally, by screening breeding pairs that produced offspring of both color morphs, we were able to exclude these two mutations as potential causative factors in one of the pedigrees (table S10). Thus, the lead variant from the association mapping analyses emerges as the sole candidate causal mutation to explain the color polymorphism. The mismatched genotype in one individual may be attributable to genetic heterogeneity or epistatic interactions with unknown genetic factors located elsewhere in the genome.

We then hypothesized that the candidate non-coding variant might be involved in the regulation of *ALDH3A2* expression as a result of its occurrence within a cis-regulatory element (i.e., enhancer/promoter). To test this hypothesis, we assessed genome-wide chromatin accessibility in regenerating feather follicles of budgerigars using a single-nucleus transposase-accessible chromatin sequencing assay (snATAC-seq, n = 2; table S9). This method identifies open chromatin regions expected to be enriched for regulatory elements. An annotation based on open chromatin profiles at the promoters of cell type markers in 1700 nuclei validated the existence of the primary cell types previously identified through scRNA-seq (Fig. 5A and fig. S13) (*34*). The only difference compared with the scRNA-seq data are that leukocytes are now collapsed in a single cluster. Among keratinocytes (defined by openness at the promoter region of several keratin genes, including *KRT17L*; Fig. 5B and fig. S14), we observed significant differential accessibility at the promoter regions of differentiation genes such as *SCEL* and *EDQM3* in cells likely corresponding to late-differentiating keratinocytes as identified in our scRNA-seq analyses (*SCEL*: $Log_2FC = 2.7$, $P = 3 \times 10^{-29}$; *EDQM3*: $Log_2FC = 3.6$, $P = 2 \times 10^{-26}$; Fig. 5C and fig. S14). Although the promoter region of *ALDH3A2* is broadly open across multiple cell types—as expected given this gene's ubiquitous expression—there is an additional open chromatin region upstream of the promoter that is specific to keratinocytes (Fig. 5D). Furthermore, there is another region of open chromatin (budgerigar genome, chr13:8,487,599-8,488,513) immediately downstream of *ALDH3A2* that is specific to the late-differentiating keratinocyte cluster ($Log_2FC = 3.40$, $P = 5.97 \times 10^{-14}$; Fig. 5, C and D). Most notably, the homologous region in the dusky lory genome includes our candidate causal variant.

To further investigate whether the region that contains the candidate causal variant might be capable of cis-regulatory activity in keratinocytes, we conducted an enrichment analysis of transcription factor (TF) binding sites within the accessible chromatin regions of the regenerating feather follicles (table S11). Several of these motifs appeared to contribute to chromatin openness in the region harboring the causal variant, as revealed by modeling the higher-order syntax of TF binding motifs of late differentiating keratinocytes using a convolutional neural network (Fig. 5E). The same method also predicted similar chromatin accessibility profiles for the *ALDH3A2* locus in the budgerigar, dusky lory, and several other parrots (fig. S15). For example, 33 bp upstream of the candidate mutation, we identified a consensus binding site motif for the RUNX family of pioneer transcription factors with a strong predicted contribution to chromatin accessibility (Fig. 5E). Since none of the motifs enriched in late-differentiating keratinocytes overlapped with the candidate causal variant, we used a comprehensive collection of TF binding motifs to search for motifs that do overlap and compiled a list of those that show differential predicted binding affinity between the T (*yellow*) and C (*red*) dusky lory alleles (Figs. 5F, fig. S16, and data S2). This catalog provides a manageable list of candidate TFs for future functional testing.

We next investigated evolutionary conservation at individual sites across the candidate region. Alignment of the genomes of 363 diverse bird species revealed only low-to-moderate phylogenetic conservation in the region (Fig. 5G and fig. S15), with the nucleotide position homologous to the candidate causative mutation in the dusky lory (budgerigar: chr13:8,487,876; dusky lory: scaffold_13:6,288,712) corresponding to a local conservation maximum. Importantly, the alignment of this region among 100 parrot genomes revealed universal conservation of cytosine at the candidate causative position and strong conservation in the flanking nucleotides (Fig. 5G). This observation suggests that strong purifying selection and ancestral functional constraints have acted upon the locus throughout parrot evolution, a pattern consistent with the preservation of a TF binding site. Collectively, our results suggest a mechanism whereby a point mutation alters the binding of a yet-unidentified TF within a cell type–specific enhancer in parrots. This alteration is likely to lead to changes in allelic expression of *ALDH3A2* during feather development in the dusky lory.

The enzyme encoded by *ALDH3A2* oxidizes aldehyde psittacofulvins to carboxyl forms

Our genetic and chemical results indicate that a substantial portion of the spectrum of parrot plumage colors can be attributed to the ratio of carboxyl- to aldehyde-containing psittacofulvins deposited during feather development. To investigate the role of *ALDH3A2* in psittacofulvin biosynthesis, we used baker's yeast (*Saccharomyces cerevisiae*) to assay psittacofulvin production upon transfection with the avian polyketide synthase (*PKS*) (*26*), with and without *ALDH3A2*. We introduced a construct

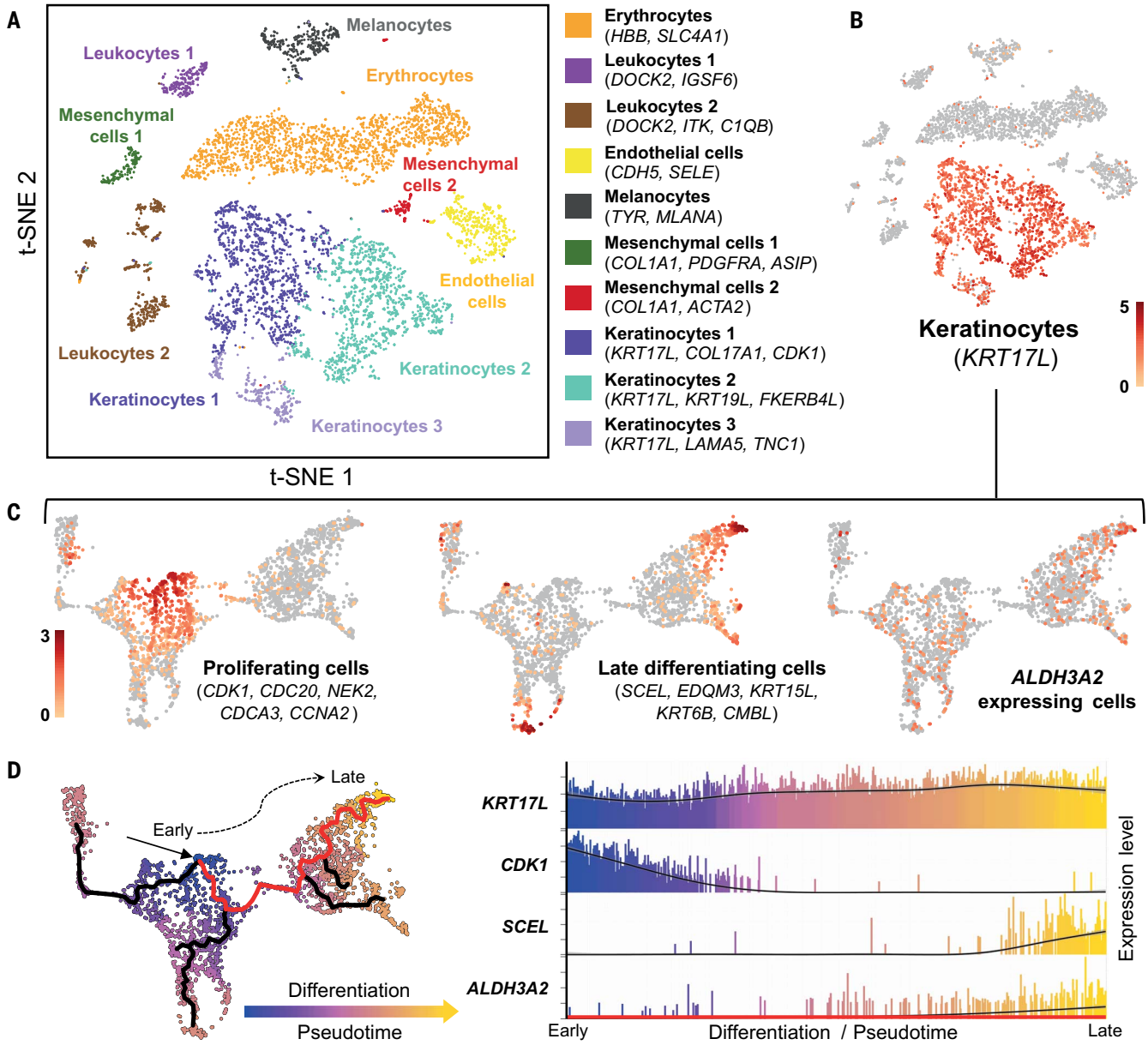

Fig. 4. *ALDH3A2* expression during feather development. (**A**) scRNA-seq analyses of budgerigar regenerating feather follicles (t-SNE projection): annotation of 6262 cells clustered by gene expression profiles into 10 major clusters. Plots of selected marker genes supporting the annotation are reported for each cluster in fig. S11. (**B**) Expression of *keratin 17-like* (*KRT17L*; ENSMUNG00000017214.1) in keratinocyte clusters [t-distributed stochastic neighbor embedding (t-SNE)]. (**C**) scRNA-seq analyses of keratinocytes [*n* = 2753 cells; uniform manifold approximation and projection (UMAP) projections]. (Left) heatmap of average expression levels of five cell cycle genes defining a sub-cluster of dividing keratinocytes (i.e., follicle proliferation zone). (Middle) heat map of average expression levels of five genes defining late-differentiating keratinocytes (fig. S12) (*34*). (Right) heatmap of *ALDH3A2* expression. (**D**) Analyses of keratinocyte differentiation. (Left) branching trajectory reflecting differentiation from dividing cells in the proliferative zone towards cells forming specialized structures in the follicle (i.e., the marginal, axial, and barbule plates). The color indicates the distance of each cell (pseudotime) from the root node in the proliferative zone (solid arrow); blue indicates early cells and yellow indicates late-differentiating cells. The red line shows the trajectory leading to a subpopulation of keratinocytes with the highest expression of *ALDH3A2*, likely axial plate cells (*34*). (Right) normalized gene expression of *KRT17L* (all keratinocytes), *CDK1* (proliferating keratinocytes), *SCEL* (late differentiating keratinocytes), and *ALDH3A2* in the cells along the trajectory (*n* = 859 cells). *ALDH3A2* expression is enriched towards late differentiating keratinocytes.

containing *PKS* into wild-type (WT) yeast and two genetically modified strains (Fig. 6A). In one strain (designated *Δhfd1* + *PKS*), we knocked out *HFD1*, the yeast ortholog of avian *ALDH3A2* (*44*). In the second modified strain, we replaced *HFD1* with *ALDH3A2* from the dusky lory (*Δhfd1* + *PKS* + *ALDH3A2*).

Through liquid chromatography (UHPLC) and mass spectrometry analyses (HRAM Q-TOF) of pigment extracts from the yeast strains (Fig. 6, A to C, and figs. S17 and S18), we identified three major peaks that stand out in their UV/VIS absorbance intensity and have chromatographic, spectrophotometric (*45*), and mass spectrometric (*19*, *26*) characteristics corresponding to psittacofulvins or their derivatives (Fig. 6, A and B). Peaks 1 and 3 closely resembled the pigments present in parrot feathers, and based on their maximum absorbance (402 nm and 418 nm) and molecular masses (243.1385 m/z and 227.1436 m/z), we

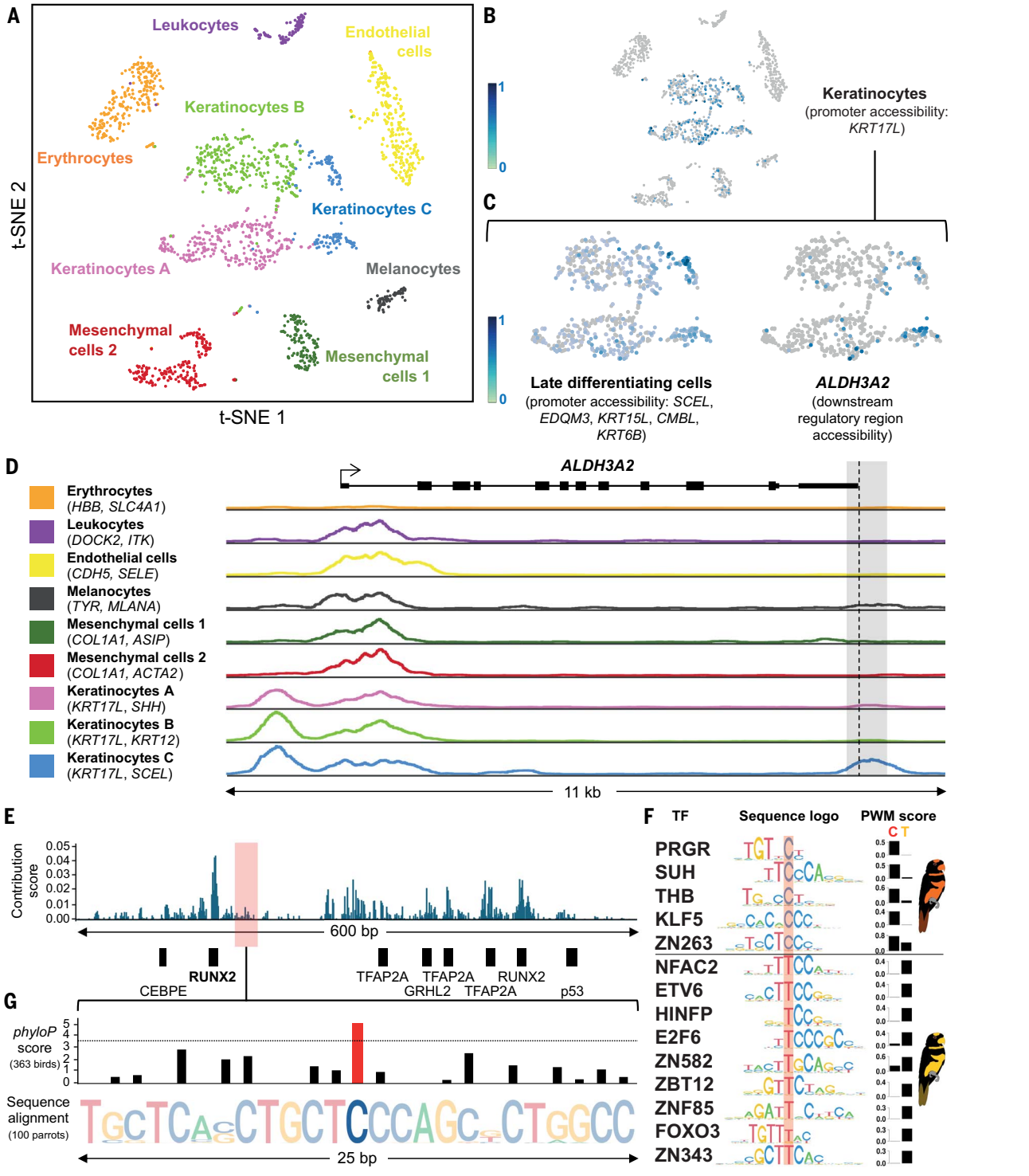

Fig. 5. A regulatory region overlaps the candidate causal mutation. (**A**) snATAC-seq analyses (t-SNE projection): annotation of 1700 cells into nine major clusters based on chromatin accessibility profiles. (**B**) Accessibility at the *Keratin 17-like* promoter in the keratinocyte clusters (t-SNE projection; gene ID: ENSMUNG00000017214.1). (**C**) snATAC-seq analyses of keratinocytes (t-SNE projections). (Left) heatmap of averaged DNA accessibility at the promoters of five genes identified in the scRNA-seq analyses as defining late differentiating keratinocytes (Fig. 4C and figs. S12 and S14) (*34*). (Right) heatmap of DNA accessibility at the ATAC peak identified downstream of *ALDH3A2* and corresponding to a late-differentiating keratinocyte–specific regulatory element. (**D**) Chromatin accessibility at the *ALDH3A2* locus for different cell types; normalized transposase cut site counts per cluster smoothed over 400-bp windows. The gray area highlights the region shown in (E). (**E**) Characterization of the regulatory element downstream of *ALDH3A2* in budgerigars. (Top) predicted nucleotide contribution for chromatin accessibility (per-nucleotide averaged contribution score from three independently trained models). (Bottom) annotation of predicted TF binding sites enriched in late-differentiating keratinocytes. The red box highlights the region shown in (G). (**F**) Sequence logos for 14 representative motifs (chosen from 10 motif subfamilies) among the top

57 predicted TF binding sites which showed the greatest change in position-weight matrix (PWM) score between C and T nucleotides (shown on the right, rounded to one significant digit). (**G**) (Top) per-nucleotide evolutionary conservation (*phyloP* scores) across 363 bird genomes projected to the budgerigar sequence. Only positive scores, indicating slower evolution than expected, are reported. The dashed line represents the non-coding genome-wide top fifth percentile. (Bottom) per-nucleotide evolutionary conservation across 100 parrot genomes. Nucleotide symbols at the same position are scaled according to their frequency. The height of the stacked symbols describes the information content at each position in the alignment.

determined them to be C16 psittacofulvins with either a carboxyl (peak 1) or an aldehyde group (peak 3). Peak 2 exhibited absorbance spectra resembling those of alcohols produced by the chemical reduction of psittacofulvins (*19*, *45*). Further analysis revealed that peak 2 splits into two closely eluting peaks, 2a and 2b (Fig. 6C), both displaying identical chromophores (Fig. 6B) and exhibiting molecular masses of 245.1537 m/z ($C_{16}H_{20}O_2$) and 231.1742 m/z ($C_{16}H_{22}O$), respectively (fig. S18). Several chemical structures are consistent with each of these molecular weights. We propose that peak 2a may arise from early chain termination before the formation of the seventh double bond. Peak 2b appears to have a terminal hydroxyl group (Fig. 6B), likely formed from the rapid conversion of the corresponding aldehyde through the action of an endogenous yeast enzyme, as previously observed for other metabolic pathways in yeast strains lacking *HFD1* (*44*).

Upon deleting *HFD1* (*Δhfd1* + *PKS*), we observed a significant alteration in pigment composition compared with the chromatogram of WT + *PKS* (Fig. 6A). This change was characterized by a noticeable decrease in the absorption intensity of peak 1 (the carboxyl form) and the concomitant increase in the absorption intensity of peaks 2 (i.e., alcohol forms) and 3 (i.e., aldehyde form). Notably, the effect of *HFD1* deletion was entirely reversed by knocking *ALDH3A2* into the locus (Fig. 6A). These results suggest that the *ALDH3A2* enzyme converts the aldehyde end group of psittacofulvins into carboxylic acid and that the orthologous yeast enzyme encoded by *HFD1* possesses the same biochemical activity.

Collectively, these findings support a model in which red aldehyde–containing psittacofulvins are the primary products of PKS (possibly released by the action of an unknown thioesterase), which are then subjected to enzymatic modification by ALDH3A2, resulting in the formation of yellow carboxyl forms (Fig. 6D). Our proposed model implies that modulation of *ALDH3A2* expression levels is necessary and sufficient to explain a large portion of the observed color variation among parrot species.

Discussion

Evolutionary innovations often act as catalysts of biological diversification. This study investigates the biochemical and genetic basis of a pigmentary system that evolved exclusively in parrots and that drives the vivid hues ornamenting their plumage. Our feather pigment analysis uncovered a strong correlation between the relative proportion of chemically distinct psittacofulvin molecules and color differences. This finding implies that psittacofulvin-based color variation—which has evolved numerous times independently in parrots—has a common chemical basis across divergent lineages of parrots.

Through a combination of genetic mapping, transcriptomics, and functional experimentation, we further implicated the ALDH3A2 enzyme in psittacofulvin-driven color shifts. We show that *ALDH3A2* underlies color variation in a rare instance of a parrot species that is phenotypically variable for yellow and red plumage coloration in the wild and further show that *ALDH3A2* expression is also correlated with color differences between plumage patches in another species. The color transition in both systems thus appears to be a direct outcome of changes in aldehyde metabolism by modulation of *ALDH3A2* expression that yield chemically distinct psittacofulvin pigments, which is consistent with the expectations based on our analysis of pigment composition. Although multiple genetic factors likely determine the overall color phenotype of a parrot species, our results show that substantial color shifts in psittacofulvin pigmentation can be accomplished through subtle changes in enzymatic activity. This simplicity may explain why evolutionary transitions from yellow/green to red hues and vice versa are so common in the parrot lineage.

ALDH3A2 is also notable for its role in vital cellular functions. As demonstrated by our yeast experiments, *ALDH3A2* is deeply conserved across the tree of life (*38*, *44*, *46*). Our findings therefore provide an example of an ancient gene being co-opted for a new function, which was likely enabled by cis-regulatory changes exerting their effects on specific cell types and thereby minimizing potential pleiotropic consequences. Considering that the spectrum of hues observed across parrot species appears often to be attributable to the selective deposition of carboxyl- versus aldehyde-containing psittacofulvins in feathers, it is highly plausible that ALDH3A2, or other enzymes with aldehyde dehydrogenase activity, are involved in such color differences in a wide range of parrot species.

The adaptive significance of parrot colors remains poorly understood (*12*), despite associations with predation risk, oxidative stress, feather degradation, and condition signaling in mate choice (*14*–*18*, *47*). Our findings, together with what is known about the role of aldehyde dehydrogenases in the detoxification of reactive compounds that accumulate within cells from lipid metabolism or dietary sources (*48*, *49*), lead to several hypotheses regarding the potential signaling information conveyed by psittacofulvin-based vibrant colors. For example, these colors could potentially indicate an individual's detoxification ability or state, as previously suggested for red carotenoid coloration (*48*) or could serve as indicators of physiological performance through their association with vital metabolic processes and ancient pathways for lipid metabolism (*10*). However, these explanations likely fail to account for the existence of the dusky lory color polymorphism. In fact, coexisting color morphs in nature are extremely rare among parrots. Although the ecological drivers behind this polymorphism remain largely unknown, polymorphisms in other highly gregarious species in several taxonomic groups are often linked to social status signaling and other behavioral differences, or exploitation of alternative ecological niches (*49*, *50*). Overall, this study provides insights into the origins of evolutionary novelties and opens avenues for investigating the adaptive significance of color displays more broadly.

Materials and methods

Additional materials and methods can be found in the supplementary materials.

Chemical analyses of parrot feathers

For in situ vibrational spectroscopy, parrot feathers were measured using a confocal Raman microscope (WITec alpha300 RSA, Oxford Instruments). Spectra were treated using WITec Project FIVE Plus software (Oxford Instruments). Statistical analyses of the Raman spectroscopy data were carried out using multivariate procedures based on singular value decomposition (SVD), as detailed previously (*34*). Psittacofulvin pigments were extracted from feathers using a pyridine-based protocol. Characteristic absorbance spectra of pigments were determined using a UHPLC system (Dionex Ultimate 3000, Thermo Fisher Scientific), coupled with a photodiode array detector (PDA). The identity of psittacofulvins was confirmed by HRAM Q-TOF mass spectrometry (IMPACT II, Bruker Daltonik). Data acquisition and processing were performed using *oTof Control*, *HyStar*, and *DataAnalysis* software (Bruker Daltonik). To compare differences in psittacofulvin content among feathers of varying

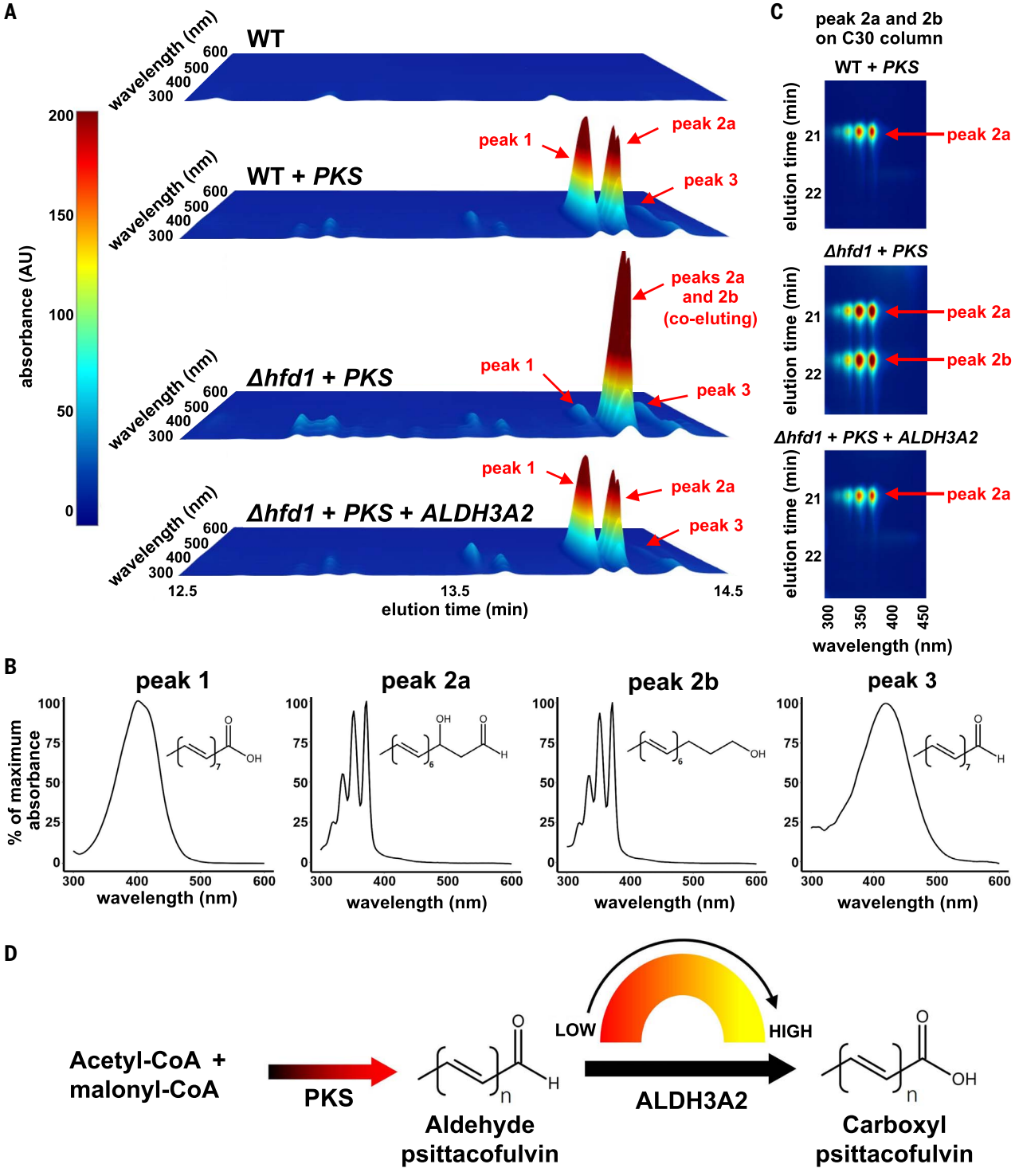

Fig. 6. The role of aldehyde dehydrogenase activity in psittacofulvin biosynthesis. (**A**) to (**C**) Analyses of yeast pigment extracts. A WT yeast strain was transformed to express *PKS* (WT + *PKS*). Two additional strains expressing *PKS* were engineered by knocking out *HFD1*, the yeast homologous of *ALDH3A2* (*Δhfd1* + *PKS*), and by knocking out *HFD1* and knocking in the dusky lory *ALDH3A2* (*Δhfd1* + *PKS* + *ALDH3A2*). (A) UHPLC spectra of yeast pigment extracts. Extracts for all the *PKS*-expressing strains contained varying amounts of chemically distinct psittacofulvins represented by three main absorbance peaks (1 to 3). Expression of *ALDH3A2* (strain: *Δhfd1* + *PKS* + *ALDH3A2*) restored the WT chromatogram (WT + *PKS*). (B) Absorption spectra of the main UHPLC peaks: peak 1 (carboxyl psittacofulvin), peaks 2a and 2b (alcohol psittacofulvins), and peak 3 (aldehyde psittacofulvin). (C) Chromatographic separation of peaks 2a and 2b. (**D**) Proposed model of psittacofulvin biosynthesis. After priming with an acetyl unit, PKS acts cyclically by adding malonyl units to extend the polyketide chain, which is then reductively released as an aldehyde. Aldehyde psittacofulvin products are then converted into the carboxyl form by ALDH3A2. Tuning of ALDH3A2 enzymatic activity (e.g., by modulation of its expression) from "low" to "high" is sufficient to explain the production of red-to-yellow psittacofulvins in parrots.

colors, a statistical analysis of our UHPLC-PDA/HRAM-QTOF data was conducted using linear mixed models in R (*51*).

Genetic mapping in the dusky lory

A draft reference genome sequence of the dusky lory was produced using PacBio HiFi sequencing reads. Contig assembly was carried out with Hifiasm (*52*, *53*) and contigs were scaffolded into a pseudochromosome assembly using the homology-based scaffolding algorithm of RagTag (*54*) and the budgerigar *bMelUnd1* genome assembly as reference (GCF_012275295.1). Gene annotation was performed using iterative runs of MAKER (*55*) following Card *et al.* (*56*). Genome-wide polymorphism data for genetic mapping was generated by whole-genome Illumina sequencing of 57 dusky lories (35 red and 22 yellow). The sequencing reads were mapped to the dusky lory draft genome assembly with BWA-MEM (*57*) and SNP and small indel variants were identified using GATK (*58*). The software Beagle (*59*, *60*) was used to impute missing genotypes and to phase the genotype data, and association analyses were run with GEMMA (*61*).

Transcriptomics

Illumina strand–specific RNA-seq libraries of regenerating feather follicles of dusky lories were generated from six samples: three red and three yellow. Trimmed reads were mapped to the dusky lory draft reference genome using HISAT2 (*62*). Differential expression analyses were carried out using DESeq2 (*63*) and Rsubread (*64*, *65*). Isoform Sequencing (Iso-seq) data using the PacBIO Iso-seq protocol was generated for regenerating feather follicles from the same six individuals. The generated HiFi reads were processed using the IsoSeq (https://isoseq.how/) lima tool and were mapped to the dusky lory draft reference genome using Minimap2 (*66*). The relative expression of the dusky lory *red* and *yellow ALDH3A2* alleles in heterozygous individuals was obtained by counting the occurrence of each transcript type in each individual. RNA-seq data of regenerating feather follicles was generated for eight rosy-faced lovebirds (*Agapornis roseicollis*) obtained from red forehead feather follicles and three additional body regions expressing green color (i.e., yellow psittacofulvin-containing patches): back feathers, chest feathers, and head feathers. Trimmed reads for each RNA-seq library were aligned to a chromosome-level *A. roseicollis* genome assembly using STAR (*67*). RSEM (*68*) was used to quantify transcript abundances from aligned RNA-seq data and edgeR (*69*) was used to normalize count data across samples using the default weighted trimmed mean of M-values (TMM) method. *ALDH3A2* expression was further validated by qPCR.

scRNA-seq and snATAC-seq

scRNA-seq and snATAC-seq experiments were conducted on regenerating feather follicles of budgerigars. Following tissue dissociation, the resuspended cells and nuclei were barcoded using a 10× Genomics Chromium instrument. The libraries were sequenced on an Illumina instrument and demultiplexed using the Cell-Ranger Fastq pipeline (10× Genomics).

The scRNA-seq libraries (n = 2) were preprocessed using the count pipeline in Cell Ranger (10× Genomics). The reads were aligned to the budgerigar genome (assembly bMelUnd1.mat.Z; Ensembl annotation release 108) using a method based on the splicing-aware aligner STAR (*67*). Dimensionality reduction of the gene expression matrix of the aggregated libraries was performed as implemented in the reanalyze function in Cell Ranger. Second-level clustering and differentiation trajectory analyses of keratinocytes were performed using Monocle 3 (*70*, *71*). Independent subclustering of keratinocytes was performed with Cell Ranger by increasing k-means clustering resolution. Differential gene expression analyses between clusters were conducted using an exact negative binomial test using the publicly available software Loupe Browser (10× Genomics). Cluster annotation was performed by inspecting each cluster's top differentially expressed genes and visual inspection of uniform manifold approximation and projection (UMAP) and t-distributed stochastic neighbor embedding (t-SNE) plots.

The snATAC-seq libraries (n = 2) were preprocessed using the count pipeline in Cell Ranger ATAC (10X Genomics) (*72*). The reads were aligned to the budgerigar genome using a method based on the BWA-MEM algorithm implemented in Cell Ranger ATAC. Normalization and dimensionality reduction of the peak-barcode matrix of the aggregated libraries were performed as implemented in the reanalyze function in Cell Ranger. Annotation of the clusters was performed using Loupe Browser. To identify cluster-specific marker genes, the analysis was focused on differential accessibility analyses within called peaks at promoter regions. A convolutional neural network method, ChromBPNet (https://github.com/kundajelab/chrombpnet), was utilized to explain the relationship between genomic sequence and base-resolution Tn5 transposase cut sites in snATAC-seq empirical data and to predict chromatin accessibility of the late differentiating keratinocytes in the dusky lory genome. The cell type–specific enriched transcription factor binding motifs were discovered de novo by HOMER (*73*).

Identification of transcription factor binding sites

Potential transcription factor binding sites overlapping the causal variant were identified using the scan_sequences function from the universalmotif library (http://bioconductor.org/packages/release/bioc/html/universalmotif.html) in R (v4.1.0). Motifs with differential predicted transcription factor binding affinity between dusky lory alleles (C versus T) were grouped into 10 sub-families of related motifs with the PWMEnrich package, R (v4.30.0); (http://bioconductor.org/packages/release/bioc/html/PWMEnrich.html).

Sequence conservation analyses

Evolutionary conservation analyses at the *ALDH3A2* locus among birds were conducted by calculating per-nucleotide phyloP scores (*74*). The scores were generated using the halPhyloP wrapper included in the whole genome aligner progressive cactus (*75*). The phyloP wrapper was run on a previously generated HAL genome alignment file incorporating sequences from 363 avian species (*76*). Conservation analyses among parrots were conducted by extracting homologous sequences surrounding the candidate causal variant in the dusky lory from 100 parrot genomes and aligning them using BioEdit (*77*).

Biochemical assays in yeast

The yeast strains used in the experiments were derived from *S. cerevisiae* strain BJ5464-NpgA. The yeast gene *HFD1* (SGD: S000004716), orthologous to the mammalian *ALDH3A2* (*44*), was knocked out from the WT strain by homologous recombination. With the same procedure, a constitutively expressed dusky lory *ALDH3A2* construct was inserted into the *HFD1* locus in the WT strain, thereby simultaneously knocking out *HFD1* and knocking in the dusky lory *ALDH3A2* sequence. The resulting yeast strains were then transformed with an expression plasmid carrying an inducible chicken polyketide synthase [*PKS*; LOC420486 (*Gallus gallus*) (*26*)] construct. The yeast strains and respective controls were cultured and the cells were harvested by centrifugation. The pigment extraction protocol was modified from the procedure described in Cooke *et al.* (*26*). The resulting extracts were subjected to HPLC and UHPLC/HRAM-QTOF analyses.

REFERENCES AND NOTES

1. I. C. Cuthill *et al.*, The biology of color. *Science* **357**, eaan0221 (2017). doi: 10.1126/science.aan0221; pmid: 28774901
2. W. Svensson, Wong, Carotenoid-based signals in behavioural ecology: A review. *Behaviour* **148**, 131–189 (2011). doi: 10.1163/000579510X548673
3. T. Amundsen, Why are female birds ornamented? *Trends Ecol. Evol.* **15**, 149–155 (2000). doi: 10.1016/S0169-5347(99)01800-5; pmid: 10717684
4. G. E. Hill, "Female choice for ornamental coloration" in *Bird Coloration, Volume 2: Function and Evolution*, G. E. Hill, K. J. McGraw, Eds. (Harvard Univ. Press, 2006), pp. 137–200.
5. R. S. Terrill, A. J. Shultz, Feather function and the evolution of birds. *Biol. Rev.* **98**, 540–566 (2023). doi: 10.1111/brv.12918; pmid: 36424880
6. N. A. Mason, R. C. K. Bowie, Plumage patterns: Ecological functions, evolutionary origins, and advances in quantification. *Auk* **137**, ukaa060 (2020). doi: 10.1093/auk/ukaa060

7. P. O. Dunn, J. K. Armenta, L. A. Whittingham, Natural and sexual selection act on different axes of variation in avian plumage color. *Sci. Adv.* **1**, e1400155 (2015). doi: 10.1126/sciadv.1400155; pmid: 26601146
8. H. M. Fox, G. Vevers, *The Nature of Animal Colors* (Macmillan, 1960).
9. R. D. H. Barrett *et al*., Linking a mutation to survival in wild mice. *Science* **363**, 499–504 (2019). doi: 10.1126/science.aav3824; pmid: 30705186
10. R. J. Weaver, R. E. Koch, G. E. Hill, What maintains signal honesty in animal colour displays used in mate choice? *Philos. Trans. R. Soc. B.* **372**, 20160343 (2017). doi: 10.1098/rstb.2016.0343; pmid: 28533460
11. A. Nemesio, Color production and evolution in parrots. *International Journal of Ornithology* **4**, 75–102 (2001).
12. M. L. Berg, A. T. D. Bennett, The evolution of plumage colouration in parrots: a review. *Emu - Austral Ornithology* **110**, 10–20 (2010). doi: 10.1071/MU09076
13. K. E. Arnold, I. P. F. Owens, N. J. Marshall, Fluorescent signaling in parrots. *Science* **295**, 92–92 (2002). doi: 10.1126/science.295.5552.92; pmid: 11778040
14. J. F. Masello, M. L. Pagnossin, T. Lubjuhn, P. Quillfeldt, Ornamental non-carotenoid red feathers of wild burrowing parrots. *Ecol. Res.* **19**, 421–432 (2004). doi: 10.1111/j.1440-1703.2004.00653.x
15. J. F. Masello, P. Quillfeldt, Body size, body condition and ornamental feathers of Burrowing Parrots: variation between years and sexes, assortative mating and influences on breeding success. *Emu - Austral Ornithology* **103**, 149–161 (2003). doi: 10.1098/rsbl.2010.0716; pmid: 20926430
16. E. H. Burtt Jr., M. R. Schroeder, L. A. Smith, J. E. Sroka, K. J. McGraw, Colourful parrot feathers resist bacterial degradation. *Biol. Lett.* **7**, 214–216 (2011). doi: 10.1098/rsbl.2010.0716; pmid: 20926430
17. L. Carballo, K. Delhey, M. Valcu, B. Kempenaers, Body size and climate as predictors of plumage colouration and sexual dichromatism in parrots. *J. Evol. Biol.* **33**, 1543–1557 (2020). doi: 10.1111/jeb.13690; pmid: 32797649
18. R. Heinsohn, S. Legge, J. A. Endler, Extreme reversed sexual dichromatism in a bird without sex role reversal. *Science* **309**, 617–619 (2005). doi: 10.1126/science.1112774; pmid: 16040708
19. R. Stradi, E. Pini, G. Celentano, The chemical structure of the pigments in *Ara macao* plumage. *CBPB* **130**, 57–63 (2001). doi: 10.1016/S1096-4959(01)00402-X; pmid: 11470444
20. C. F. W. Krukenberg, Die Federfarbstoffe der Psittaciden. *Vergleichend-physiologische Studien Reihe 2. Abtlg* **2**, 29–36 (1882).
21. M. Veronelli, G. Zerbi, R. Stradi, In situ resonance Raman spectra of carotenoids in bird's feathers. *J. Raman Spectrosc.* **26**, 683–692 (1995). doi: 10.1002/jrs.1250260815
22. K. J. McGraw, M. C. Nogare, Distribution of unique red feather pigments in parrots. *Biol. Lett.* **1**, 38–43 (2005). doi: 10.1098/rsbl.2004.0269; pmid: 17148123
23. M. B. Toomey *et al*., High-density lipoprotein receptor *SCARB1* is required for carotenoid coloration in birds. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5219–5224 (2017). doi: 10.1073/pnas.1700751114; pmid: 28465440
24. M. B. Toomey *et al*., A mechanism for red coloration in vertebrates. *Curr. Biol.* **32**, 4201–4214.e12 (2022). doi: 10.1016/j.cub.2022.08.013; pmid: 36049480
25. M. A. Gazda *et al*., A genetic mechanism for sexual dichromatism in birds. *Science* **368**, 1270–1274 (2020). doi: 10.1126/science.aba0803; pmid: 32527835
26. T. F. Cooke *et al*., Genetic Mapping and Biochemical Basis of Yellow Feather Pigmentation in Budgerigars. *Cell* **171**, 427–439.e21 (2017). doi: 10.1016/j.cell.2017.08.016; pmid: 28985565
27. N. I. Mundy, Colouration Genetics: Pretty Polymorphic Parrots. *Curr. Biol.* **28**, R113–R114 (2018). doi: 10.1016/j.cub.2017.12.045; pmid: 29408256
28. F. Ke *et al*., Convergent evolution of parrot plumage coloration. *Proc. Natl. Acad. Sci. U.S.A. Nexus* **3**, pgae107 (2024). doi: 10.1093/pnasnexus/pgae107; pmid: 38528953
29. J. E. Barnsley, E. J. Tay, K. C. Gordon, D. B. Thomas, Frequency dispersion reveals chromophore diversity and colour-tuning mechanism in parrot feathers. *R. Soc. Open Sci.* **5**, 172010 (2018). doi: 10.1098/rsos.172010; pmid: 30109049
30. T. F. Wright *et al*., A multilocus molecular phylogeny of the parrots (Psittaciformes): Support for a Gondwanan origin during the cretaceous. *Mol. Biol. Evol.* **25**, 2141–2156 (2008). doi: 10.1093/molbev/msn160; pmid: 18653733
31. E. J. Tay, J. E. Barnsley, D. B. Thomas, K. C. Gordon, Elucidating the resonance Raman spectra of psittacofulvins. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **262**, 120146 (2021). doi: 10.1016/j.saa.2021.120146; pmid: 34274684
32. M. A. Kane, J. L. Napoli, *Quantification of Endogenous Retinoids* (Springer, 2010). pp. 1–54.
33. N. Collar, P. F. D. Boesman, "Dusky Lory (Pseudeos fuscata)" in *Birds of the World* (Cornell Lab of Ornithology, 2020).
34. Materials and methods are available as supplementary materials.
35. M. Otsuka *et al*., A human transporter protein that mediates the final excretion step for toxic organic cations. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17923–17928 (2005). doi: 10.1073/pnas.0506483102; pmid: 16330770
36. T. L. Kelson, J. R. Secor McVoy, W. B. Rizzo, Human liver fatty aldehyde dehydrogenase: Microsomal localization, purification, and biochemical characterization. *Biochim. Biophys. Acta* **1335**, 99–110 (1997). doi: 10.1016/S0304-4165(96)00126-2; pmid: 9133646
37. C.-H. Chang *et al*., Sculpting skin appendages out of epidermal layers via temporally and spatially regulated apoptotic events. *J. Invest. Dermatol.* **122**, 1348–1355 (2004). doi: 10.1111/j.0022-202X.2004.22611.x; pmid: 15175023
38. G. R. Rogers, N. G. Markova, V. De Laurenzi, W. B. Rizzo, J. G. Compton, Genomic organization and expression of the human fatty aldehyde dehydrogenase gene (FALDH). *Genomics* **39**, 127–135 (1997). doi: 10.1006/geno.1996.4501; pmid: 9027499
39. Z. Yue, T. X. Jiang, P. Wu, R. B. Widelitz, C. M. Chuong, Sprouty/FGF signaling regulates the proximal-distal feather morphology and the size of dermal papillae. *Dev. Biol.* **372**, 45–54 (2012). doi: 10.1016/j.ydbio.2012.09.004; pmid: 23000358
40. M. Yu, P. Wu, R. B. Widelitz, C.-M. Chuong, The morphogenesis of feathers. *Nature* **420**, 308–312 (2002). doi: 10.1038/nature01196; pmid: 12442169
41. M.-F. Champliaud, R. E. Burgeson, W. Jin, H. P. Baden, P. F. Olson, cDNA cloning and characterization of sciellin, a LIM domain protein of the keratinocyte cornified envelope. *J. Biol. Chem.* **273**, 31547–31554 (1998). doi: 10.1074/jbc.273.47.31547; pmid: 9813070
42. B. Strasser *et al*., Evolutionary origin and diversification of epidermal barrier proteins in amniotes. *Mol. Biol. Evol.* **31**, 3194–3205 (2014). doi: 10.1093/molbev/msu251; pmid: 25169930
43. C. M. Chuong, G. M. Edelman, Expression of cell-adhesion molecules in embryonic induction. II. Morphogenesis of adult feathers. *J. Cell Biol.* **101**, 1027–1043 (1985). doi: 10.1083/jcb.101.3.1027; pmid: 3897242
44. K. Nakahara *et al*., The Sjögren-Larsson syndrome gene encodes a hexadecenal dehydrogenase of the sphingosine 1-phosphate degradation pathway. *Mol. Cell* **46**, 461–471 (2012). doi: 10.1016/j.molcel.2012.04.033; pmid: 22633490
45. F. Adamec *et al*., Spectroscopic investigation of a brightly colored psittacofulvin pigment from parrot feathers. *Chem. Phys. Lett.* **648**, 195–199 (2016). doi: 10.1016/j.cplett.2016.02.029
46. Z. Lin, G. Carney, W. B. Rizzo, Genomic organization, expression, and alternate splicing of the mouse fatty aldehyde dehydrogenase gene. *Mol. Genet. Metab.* **71**, 496–505 (2000). doi: 10.1006/mgme.2000.3084; pmid: 11073717
47. R. Morelli, R. Loscalzo, R. Stradi, A. Bertelli, M. Falchi, Evaluation of the antioxidant activity of new carotenoid-like compounds by electron paramagnetic resonance. *Drugs Exp. Clin. Res.* **29**, 95–100 (2003). pmid: 14708454
48. N. I. I. Mundy *et al*., Red carotenoid coloration in the zebra finch is controlled by a cytochrome P450 gene cluster. *Curr. Biol.* **26**, 1435–1440 (2016). doi: 10.1016/j.cub.2016.04.047; pmid: 27212402
49. A. Roulin, The evolution, maintenance and adaptive function of genetic colour polymorphism in birds. *Biol. Rev.* **79**, 815–848 (2004). doi: 10.1017/S1464793104006487; pmid: 15682872
50. S. R. Pryke, Fiery red heads: Female dominance among head color morphs in the Gouldian finch. *Behav. Ecol.* **18**, 621–627 (2007). doi: 10.1093/beheco/arm020
51. R Core Team, R: A language and environment for statistical computing (R Foundation for Statistical Computing, 2022); www.R-project.org/.
52. H. Cheng *et al*., Haplotype-resolved assembly of diploid genomes without parental data. *Nat. Biotechnol.* **40**, 1332–1335 (2022). doi: 10.1038/s41587-022-01261-x; pmid: 35332338
53. 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). doi: 10.1038/s41592-020-01056-5; pmid: 33526886
54. M. Alonge *et al*., Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* **23**, 258 (2022). doi: 10.1186/s13059-022-02823-7; pmid: 36522651
55. C. Holt, M. Yandell, MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**, 491 (2011). doi: 10.1186/1471-2105-12-491; pmid: 22192575
56. D. C. Card *et al*., Genomic basis of convergent island phenotypes in Boa constrictors. *Genome Biol. Evol.* **11**, 3123–3143 (2019). doi: 10.1093/gbe/evz226; pmid: 31642474
57. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). doi: 10.1093/bioinformatics/btp324; pmid: 19451168
58. A. McKenna *et al*., The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010). doi: 10.1101/gr.107524.110; pmid: 20644199
59. 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). doi: 10.1016/j.ajhg.2018.07.015; pmid: 30100085
60. B. L. Browning, X. Tian, Y. Zhou, S. R. Browning, Fast two-stage phasing of large-scale sequence data. *Am. J. Hum. Genet.* **108**, 1880–1890 (2021). doi: 10.1016/j.ajhg.2021.08.005; pmid: 34478634
61. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012). doi: 10.1038/ng.2310; pmid: 22706312
62. D. Kim, B. Langmead, S. L. Salzberg, HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015). doi: 10.1038/nmeth.3317; pmid: 25751142
63. 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). doi: 10.1186/s13059-014-0550-8; pmid: 25516281
64. 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). doi: 10.1093/bioinformatics/btt656; pmid: 24227677
65. 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 (2019). doi: 10.1093/nar/gkz114; pmid: 30783653
66. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018). doi: 10.1093/bioinformatics/bty191; pmid: 29750242
67. A. Dobin *et al*., STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). doi: 10.1093/bioinformatics/bts635; pmid: 23104886
68. B. Li, C. N. Dewey, RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011). doi: 10.1186/1471-2105-12-323; pmid: 21816040
69. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010). doi: 10.1093/bioinformatics/btp616; pmid: 19910308
70. J. Cao *et al*., The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019). doi: 10.1038/s41586-019-0969-x; pmid: 30787437
71. C. Trapnell *et al*., The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014). doi: 10.1038/nbt.2859; pmid: 24658644
72. A. T. Satpathy *et al*., Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat. Biotechnol.* **37**, 925–936 (2019). doi: 10.1038/s41587-019-0206-z; pmid: 31375813
73. S. Heinz *et al*., Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010). doi: 10.1016/j.molcel.2010.05.004; pmid: 20513432
74. K. S. Pollard, M. J. Hubisz, K. R. Rosenbloom, A. Siepel, Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20**, 110–121 (2010). doi: 10.1101/gr.097857.109; pmid: 19858363
75. J. Armstrong *et al*., Progressive Cactus is a multiple-genome aligner for the thousand-genome era. *Nature* **587**, 246–251 (2020). doi: 10.1038/s41586-020-2871-y; pmid: 33177663
76. E. D. Jarvis *et al*., Phylogenomic analyses data of the avian phylogenomics project. *Gigascience* **4**, 4 (2015). doi: 10.1186/s13742-014-0038-1; pmid: 25741440

77. T. A. Hall, BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98 (1999).
78. J. Brejcha *et al*., Data from: A molecular mechanism for bright color variation in parrots, Dryad (2024); http://dx.doi.org/10.7927/H4M61H5F.

ACKNOWLEDGMENTS

We thank T. Cooke for his advice on psittacofulvin biosynthetic assays. We thank the following breeders who provided samples and feathers for this study: C. F. Macedo, R. Sousa, G. George, R. Clemente, R. Vašata, M. Hřebík from Pilsen Zoo, M. Žohová from Laguna rescue center, and L. Žoha from KPEP. 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 supplies. We thank N. Da Silva for the yeast strains. We thank M. Kořínek, N. Perrins, R. Hynson, and the Macaulay Library at Cornell Lab of Ornithology for permission to use parrot photos. We thank C. Myers for advice in single-cell wet lab practices. We thank J. C. Carvalho for producing the graphical abstract. We thank the CCMR staff, E. S. N. Lo, T. S. Ngai, and M. Y. Wu for animal care and husbandry, and the research computing facilities offered by the Information Technology Services, the University of Hong Kong. **Funding:** This work was funded by the following: European Research Council under the European Union's Horizon 2020 research and innovation program; grant agreement No. 101000504 (to M.C.); Portuguese Foundation for Science and Technology (FCT, https://www.fct.pt) research fellowships SFRH/BD/147030/2019 and PD/BD/128492/2017 in the scope of the Biodiversity, Genetics and Evolution (BIODIV) PhD program (to C.I.M. and P.P.); Portuguese Foundation for Science and Technology (FCT, https://www.fct.pt) research contracts CEECINST/00014/2018/CP1512/CT0002, 2020.01405.CEECIND/CP1601/CT0011, and 2020.01494.CEECIND (to M.C., P.A., and P.M.A.); Portuguese National Funds (Transitory Norm contract [DL57/2016/CP1440/CT0006]) (to R.J.L.); Portuguese Foundation for Science and Technology (FCT, https://www.fct.pt); R&D Project in All Scientific Domains 2022.06261.PTDC. (to R.J.L.); FWO (Fonds voor Wetenschappelijk Onderzoek Vlaanderen) travel grant (to M.N.); Universiteit Gent BOF grant (to M.N.); BAEF (Belgian American Educational Foundation) fellowship (to M.N.); Seed Fund for Basic Research; University of Hong Kong (to S.J.S.); Research Infrastructure METROFOOD-CZ supported by the Ministry of Education, Youth, and Sports of the Czech Republic; project No. LM2023064 (to P.Ma.). **Author contributions:** Conceptualization: M.C., J.C.C., and P.M.A. Methodology: J.B., J.C.C., M.C., M.W., S.B., R.Ar., E.S.K.P., S.Y.W.S. Investigation: C.I.M., G.D., J.B., M.P.J.N., P.A., P.Ma., P.Mo., R.Af., R.Ar., S.A., S.B., S.J.S., Y.L., Y.O., A.C., E.S.K.P., S.Y.W.S. Visualization: C.I.M., G.D., J.B., M.J.P.N., R.Ar., S.B., and A.C. Experimental procedures on animals: R.J.L., E.S.K.P., S.Y.W.S. Sampling and animal husbandry: J.B., P.M.A., R.J.L., S.G.R., U.A., E.S.K.P., S.Y.W.S. Funding acquisition: J.C.C., M.C. Supervision: J.C.C. and M.C. Writing – original draft: J.B., J.C.C., M.C., R.Ar., S.B. Writing – review and editing: All authors. **Competing interests:** R.Ar., S.B., J.B., P.M.A., J.C.C., and M.C. are inventors on a pending patent application related to the technology described in this work. The remaining authors declare no conflict of interest. **Data and materials availability:** PacBio and Illumina reads are available in NCBI SRA under BioProject PRJNA986688. Gene annotation, the draft reference genome, and the raw Raman and HPLC data are available at Dryad (*78*).

SUPPLEMENTARY MATERIALS

Submitted 12 April 2024; accepted 5 September 2024
10.1126/science.adp7710