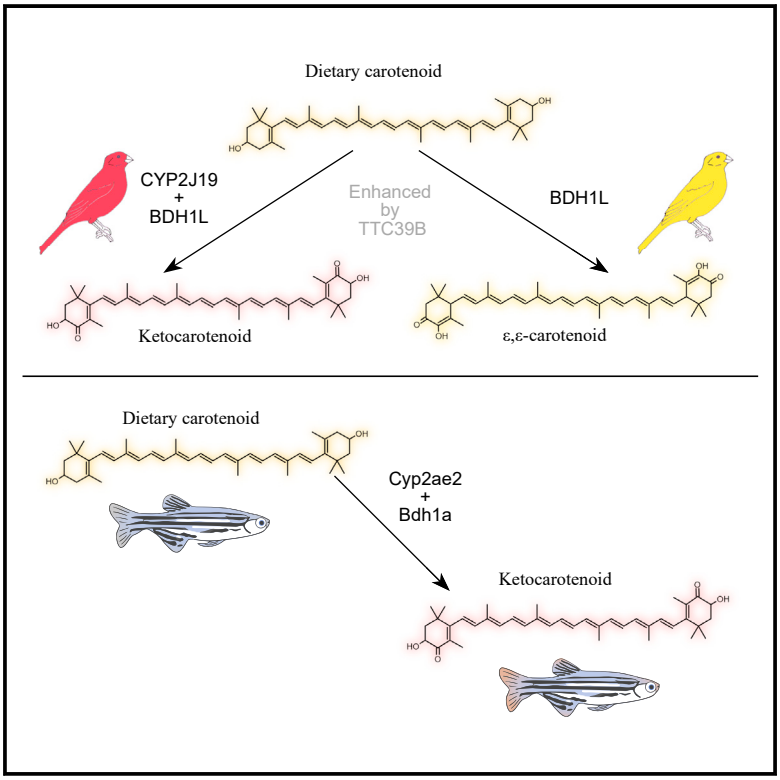
Article

Current Biology

A mechanism for red coloration in vertebrates

Graphical abstract



Highlights

- Avian CYP2J19 and BDH1L together convert dietary carotenoids into ketocarotenoids
- TTC39B enhances the production of ketocarotenoids by CYP2J19 and BDH1L
- BDH1L alone converts dietary carotenoids into ε,ε-carotenoids
- Homologs of CYP2J19 and BDH1L mediate ketocarotenoid biosynthesis in fish

Authors

Matthew B. Toomey,
Cristiana I. Marques,
Pedro M. Araújo, ..., David M. Parichy,
Miguel Carneiro, Joseph C. Corbo

Correspondence

mbt6332@utulsa.edu (M.B.T.),
miguel.carneiro@cibio.up.pt (M.C.),
jcorbo@wustl.edu (J.C.C.)

In brief

Red ketocarotenoid-based coloration is an important signal for many vertebrates. Toomey et al. describe a two-step enzymatic mechanism that mediates the production of ketocarotenoids in birds and fish. This work opens new avenues for understanding the evolution and function of red coloration in vertebrates.

Toomey et al., 2022, Current Biology *32*, 1–14
October 10, 2022
https://doi.org/10.1016/j.cub.2022.08.013

CellPress

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

Current Biology

CellPress

Article

A mechanism for red coloration in vertebrates

Matthew B. Toomey,[1,12,*] Cristiana I. Marques,[2,3,4] Pedro M. Araújo,[2,3,5] Delai Huang,[6] Siqiong Zhong,[7] Yu Liu,[8] Gretchen D. Schreiner,[8] Connie A. Myers,[8] Paulo Pereira,[2,3,4] Sandra Afonso,[2,3] Pedro Andrade,[2,3] Małgorzata A. Gazda,[2,11] Ricardo J. Lopes,[2,3,9] Ivan Viegas,[10] Rebecca E. Koch,[1] Maureen E. Haynes,[1] Dustin J. Smith,[1] Yohey Ogawa,[8] Daniel Murphy,[8] Rachel E. Kopec,[7] David M. Parichy,[6] Miguel Carneiro,[2,3,13,*] and Joseph C. Corbo[8,14,15,*]

[1]Department of Biological Science, University of Tulsa, Tulsa, OK, USA
[2]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO, Universidade do Porto, Vairão, Portugal
[3]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Vairão, Portugal
[4]Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal
[5]University of Coimbra, MARE – Marine and Environmental Sciences Centre, Department of Life Sciences, Coimbra, Portugal
[6]Department of Biology and Department of Cell Biology, University of Virginia, Charlottesville, VA, USA
[7]Program in Human Nutrition, Department of Human Sciences, Ohio State University, Columbus, OH, USA
[8]Department of Pathology and Immunology, Washington University School of Medicine, St Louis, MO, USA
[9]MHNC-UP, Natural History and Science Museum of the University of Porto, Porto, Portugal
[10]University of Coimbra, Centre for Functional Ecology, Department of Life Sciences, Coimbra, Portugal
[11]Present address: Institut Pasteur, Université de Paris, CNRS UMR 3525, INSERM UA12, Comparative Functional Genomics group, 75015 Paris, France
[12]Twitter: @mbtoomey
[13]Twitter: @evolgenCIBIO
[14]Twitter: @CorboLab
[15]Lead contact
*Correspondence: mbt6332@utulsa.edu (M.B.T.), miguel.carneiro@cibio.up.pt (M.C.), jcorbo@wustl.edu (J.C.C.)
https://doi.org/10.1016/j.cub.2022.08.013

SUMMARY

Red coloration is a salient feature of the natural world. Many vertebrates produce red color by converting dietary yellow carotenoids into red ketocarotenoids via an unknown mechanism. Here, we show that two enzymes, cytochrome P450 2J19 (CYP2J19) and 3-hydroxybutyrate dehydrogenase 1-like (BDH1L), are sufficient to catalyze this conversion. In birds, both enzymes are expressed at the sites of ketocarotenoid biosynthesis (feather follicles and red cone photoreceptors), and genetic evidence implicates these enzymes in yellow/red color variation in feathers. In fish, the homologs of CYP2J19 and BDH1L are required for ketocarotenoid production, and we show that these enzymes are sufficient to produce ketocarotenoids in cell culture and when ectopically expressed in fish skin. Finally, we demonstrate that the red-cone-enriched tetratricopeptide repeat protein 39B (TTC39B) enhances ketocarotenoid production when co-expressed with CYP2J19 and BDH1L. The discovery of this mechanism of ketocarotenoid biosynthesis has major implications for understanding the evolution of color diversity in vertebrates.

INTRODUCTION

Yellow, orange, and red colors play important and diverse roles in vertebrate signaling.[1–3] In many species, these colors are produced via deposition of carotenoid pigments in the integument.[4–6] In some clades (e.g., reptiles and birds), carotenoids are also deposited in retinal cone oil droplets,[7,8] specialized optical organelles that spectrally filter incoming light, thereby reducing spectral overlap between photoreceptor subtypes and consequently enhancing color discrimination.[9] Vertebrates cannot synthesize carotenoids *de novo* but must obtain them from their diet.[5] Diet-derived yellow carotenoids such as β-carotene, lutein, and zeaxanthin are either deposited directly in tissues or are chemically modified, then deposited.[4,10]

Three major classes of modified carotenoids are found in vertebrates: apocarotenoids, 4-keto-carotenoids (herein referred to simply as "ketocarotenoids"), and ε,ε-carotenoids (Figure 1).[5,11] Apocarotenoids such as all-*trans* retinol (vitamin A) and galloxanthin are derived from dietary yellow carotenoids via oxidative cleavage of the polyene chain.[12,13] Ketocarotenoids such as astaxanthin, canthaxanthin, and α-doradexanthin are derived from dietary yellow carotenoids via addition of ketones at the C4 position of one or both terminal β-ionone rings.[4,10,14] These ketones extend the conjugated system, thereby endowing ketocarotenoids with red-shifted absorption.[4,10] Modified ε,ε-carotenoids such as canary xanthophyll B are derived from dietary yellow carotenoids via a β-to-ε shift in the ring double bond and oxidation of the 3-hydroxyl group;[10] ε,ε-carotenoids are found in the feathers of diverse bird species where they frequently co-occur with ketocarotenoids.[15,16]

In vertebrates, apocarotenoid formation is well understood,[12,13] but relatively little is known about the mechanisms of ketocarotenoid and ε,ε-carotenoid biosynthesis.[19,20] Previous studies showed that CYP2J19 is required for ketocarotenoid production in birds,[21,22] but we found that CYP2J19 alone is not sufficient for this chemical transformation (see below). We

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

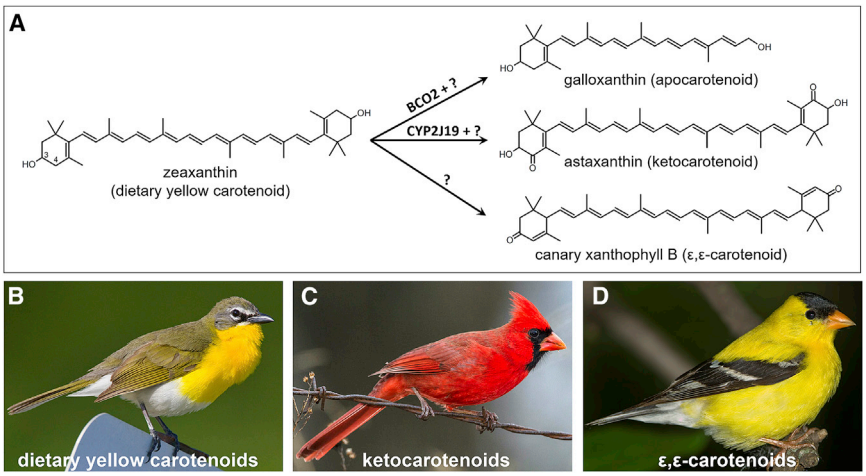


Figure 1. Three major classes of modified carotenoids in vertebrates

(A) Dietary carotenoids (e.g., zeaxanthin, lutein, and β-carotene) can be converted endogenously into three main classes of modified carotenoid: apocarotenoids, ketocarotenoids, and ε,ε-carotenoids. Here, the term "ε,ε-carotenoid" refers specifically to modified carotenoids with two ε-rings and a ketone group at C3. BCO2 cleaves zeaxanthin to produce the aldehyde form of galloxanthin, which is subsequently reduced by an unknown enzyme. CYP2J19 is necessary, but not sufficient, for addition of a ketone group at C4. The mechanism of ε,ε-carotenoid formation was unknown prior to the present work.

(B–D) The yellow-breasted chat (*Icteria virens*) (B), the northern cardinal (*Cardinalis cardinalis*) (C), and the American goldfinch (*Spinus tristis*) (D) pigment their feathers with the indicated carotenoid types.[17,18] Photo credits: Geoff Hill.

therefore searched for additional factors that might work in concert with CYP2J19 to catalyze ketocarotenoid formation. This search led to the discovery of an enzymatic pathway capable of converting dietary yellow carotenoids into red ketocarotenoids and ε,ε-carotenoids. This finding has far-reaching implications for understanding color vision and vertebrate ornamental coloration.

RESULTS

Identifying candidate factors for ketocarotenoid biosynthesis

To identify factors required for ketocarotenoid production, we used RNA-seq to profile developing and mature chicken cone photoreceptors. Five of the six avian cone subtypes accumulate carotenoids in their inner segments (Figures 2A–2D),[7,23] making them an outstanding system in which to elucidate molecular mechanisms of carotenoid modification. Yolk- or diet-derived zeaxanthin is either deposited unmodified (in green and the accessory member [DC-A] of the double cone) or modified in a subtype-specific fashion and then deposited in oil droplets (blue, red, and principal member [DC-P] of the double cone) (Figure 2D). In red cones, zeaxanthin is converted into astaxanthin.[7,8,24] We leveraged the unique patterns of autofluorescence conferred by carotenoids to sort chicken cones at hatching (P0) and at post-hatch day 15 (P15) by fluorescence-activated cell sorting (FACS) (Figures 2D–2F). We confirmed the identity of individual sorted cone subtypes by qPCR (Figure 2G; Table S1) and then subjected the sorted cones to RNA-seq analysis. We found 582 genes to be differentially expressed among cone subtypes at P0 and 938 genes at P15 (Data S1). We classified differentially expressed genes according to their patterns of expression among subtypes, identifying 57 genes that were enriched in red cones at either P0 or P15 (Figure 2H; Data S1). CYP2J19 was the most highly expressed gene among the 57 red-cone-enriched genes.

CYP2J19 and BDH1L are sufficient to produce ketocarotenoids

To determine whether CYP2J19 can convert zeaxanthin into astaxanthin, we expressed the gene in cultured mammalian cells and provided the cells with zeaxanthin as substrate. After incubating for 24 h, we analyzed carotenoid content using reverse-phase high-performance liquid chromatography-diode array detection (HPLC-DAD) and identified no astaxanthin in the cells. Instead, we observed a novel product (peak 2) with an absorption spectrum identical to that of zeaxanthin but with an earlier retention time (Figures 3A and 3B). These findings suggest that this novel carotenoid contains the same chromophore structure as zeaxanthin but is more polar and therefore likely represents an oxidized (possibly hydroxylated) form of zeaxanthin. Based on this result, we hypothesized that full conversion of zeaxanthin to astaxanthin might require another factor in addition to CYP2J19.

We therefore screened the list of red-cone-enriched genes for enzymes with oxidoreductase activity. We identified two genes encoding closely related members of the short-chain dehydrogenase/reductase family:[29] 11-*cis* retinol dehydrogenase (*RDH5*) and 3-hydroxybutyrate dehydrogenase 1-like (*BDH1L*). In vertebrates, RDH5 is usually expressed in the retinal pigment epithelium where it oxidizes 11-*cis* retinol to 11-*cis* retinal as part of the visual cycle.[30] BDH1L has not previously been studied biochemically but shows strong sequence similarity to BDH1, an enzyme that interconverts acetoacetate and 3-hydroxybutyrate, the principal ketone bodies produced during caloric restriction.[31] To evaluate these candidates, we co-transfected each gene individually with CYP2J19 into cultured cells, provided zeaxanthin as substrate, and then analyzed the reaction products by HPLC-DAD. Co-transfection of *CYP2J19* and *RDH5* failed to produce astaxanthin (data not shown). In contrast, co-expression of CYP2J19 and BDH1L produced a novel product (peak 4) with a retention time and absorption spectrum identical to those of astaxanthin (Figures 3A–3C). Mass spectrometry (both MS scan and MS/MS experiments) demonstrated that this new product has an identical molecular precursor ion and product ion fragmentation pattern to those of an authentic astaxanthin standard (Figures S1**A–S1D**). We also showed that CYP2J19 and BDH1L preferentially produce 3S,3′S-astaxanthin when provided with an equimolar mixture of all three zeaxanthin stereoisomers as substrate (Figure S2). This stereoselectivity may explain the predominance of 3S,3′S-astaxanthin in red cones.[32] Next, we

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

CellPress

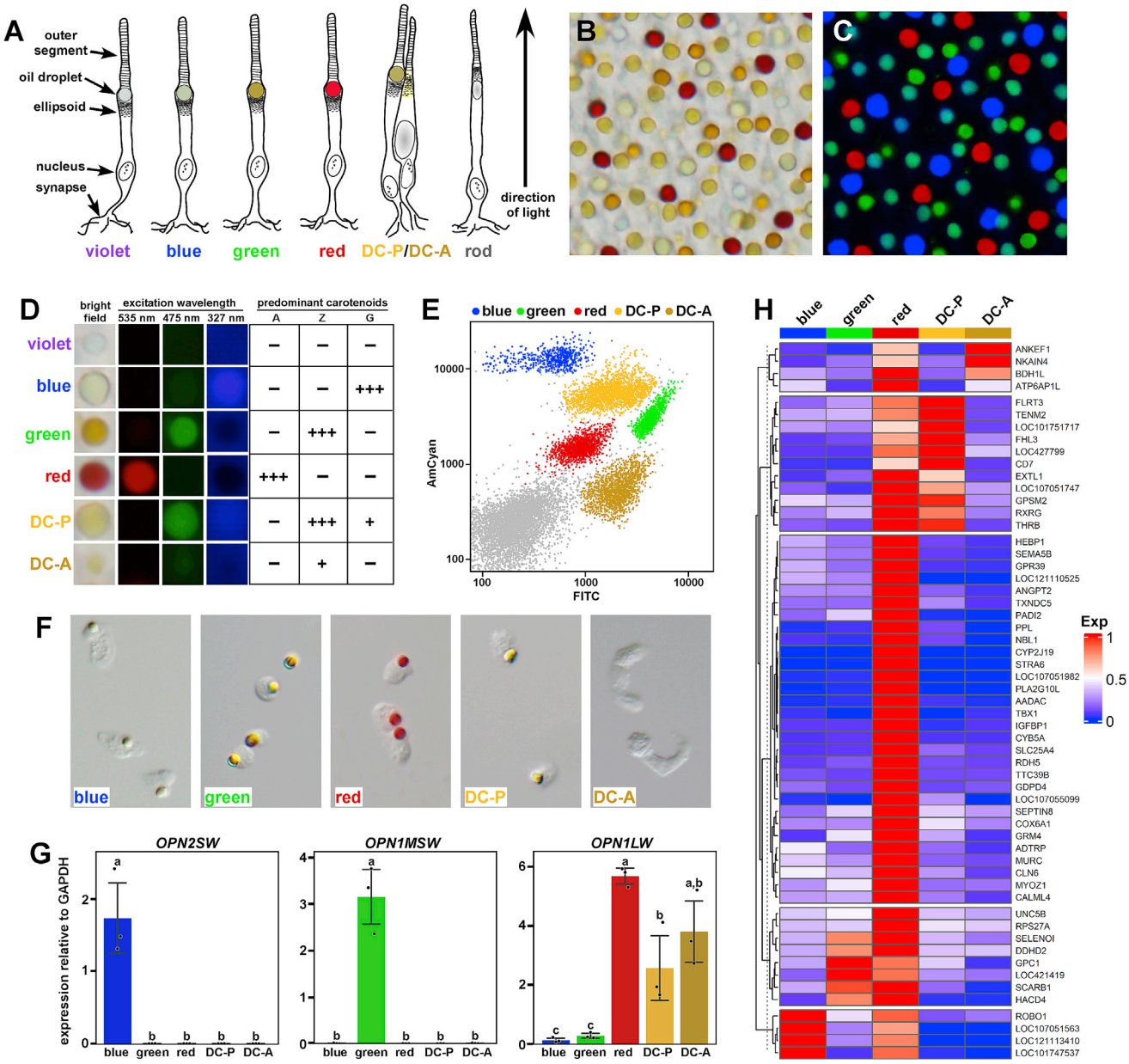


Figure 2. Oil droplet autofluorescence enables transcriptome profiling of avian cone subtypes

(A) Birds have seven photoreceptor subtypes:[25] rods, four single cones (violet, blue, green, and red), and a double cone consisting of principal (DC-P) and accessory (DC-A) members.

(B) A P15 chicken (*Gallus gallus*) retina viewed *en face*; five oil droplet types are visible.

(C) Carotenoid-containing oil droplets have distinct patterns of autofluorescence as revealed by illumination of the retina in (B) with monochromatic light of three wavelengths (327, 475, and 535 nm).[25]

(D) Autofluorescence correlates with carotenoid content of individual cone subtypes. Note that DC-A contains carotenoids in the inner segment despite the absence of an oil droplet. A, astaxanthin; Z, zeaxanthin; G, galloxanthin.

(E) Autofluorescence enables separation of five cone subtypes by FACS.

(F) Sorted cones of a given subtype consist of an oil droplet (except for DC-A) and a portion of inner segment.

(G) qPCR on cDNA derived from sorted P0 cone subtypes confirms expected patterns of opsin expression:[26] *OPN2SW*, blue cone opsin; *OPN1MSW*, green cone opsin, *OPN1LW*, red cone opsin. Data are represented as mean ± SD (n = 3); individual values are shown as dots. The transcript levels were normalized to glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) levels. qPCR results were evaluated by one-way ANOVA and Tukey's honestly significant difference (HSD) test. Distinct letters indicate statistically significant differences (p-adj < 0.05).

(H) Heatmap showing 57 genes enriched in red cones or in red cones and one additional cone subtype at either P0 or P15. Differentially expressed genes were identified using Sleuth[27] and clustered using divisive analysis (DIANA).[28]

See also Table S1 and Data S1.

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

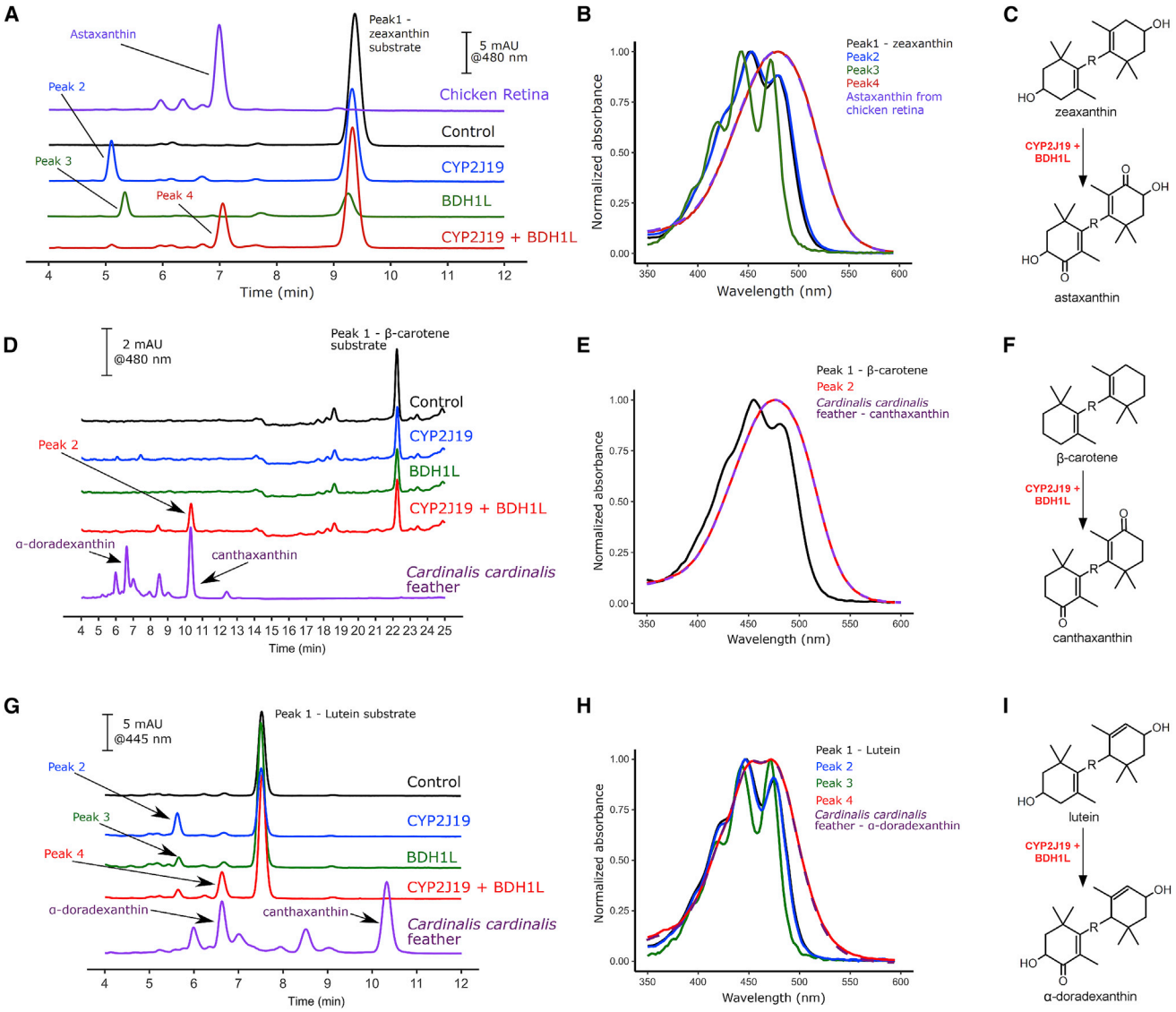


Figure 3. CYP2J19 and BDH1L are sufficient to convert dietary yellow carotenoids into ketocarotenoids

(A and B) HPLC-DAD profiles (A) and absorption spectra (B) of carotenoids extracted from HEK293 cells transfected with chicken *CYP2J19*, *BDH1L*, or both and provided with a zeaxanthin substrate. Both enzymes together convert zeaxanthin into a carotenoid with retention time and absorption spectrum identical to those of astaxanthin derived from P15 chicken retina.

(C) Proposed enzymatic mechanism based on data in (A) and (B). Carotenoid structures are truncated, and R represents the polyene chain (compare to Figure 1A).

(D and E) HPLC-DAD profiles (D) and absorbance spectra (E) of carotenoids extracted from HEK293 cells transfected with chicken *CYP2J19* and *BDH1L* and provided with a β-carotene substrate.

(F) Proposed enzymatic mechanism based on data in (D) and (E).

(G and H) HPLC-DAD profiles (G) and absorbance spectra (H) of carotenoids extracted from HEK293 cells transfected with chicken *CYP2J19* and *BDH1L* and provided with a lutein substrate.

(I) Proposed enzymatic mechanism based on data in (G) and (H).

In (D), (E), (G), and (H), the HPLC elution profile and selected spectra of carotenoids extracted from northern cardinal (*Cardinalis cardinalis*) feathers (purple traces) are shown for comparison. The scale bars in (D) and (G) apply only to the traces derived from cell extracts; the y axes of the feather sample chromatograms have been rescaled to facilitate comparison.

See also Figures S1–S3.

repeated the same cell culture assay using two other dietary yellow carotenoids (β-carotene and lutein) as substrates and observed formation of the corresponding ketocarotenoids, canthaxanthin, and α-doradexanthin, respectively (Figures 3D–3I). We therefore conclude that CYP2J19 and BDH1L together are sufficient to convert dietary yellow carotenoids into ketocarotenoids.

In the course of these experiments, we also assayed the activity of BDH1L alone on a zeaxanthin substrate. We observed production of a novel carotenoid (peak 3) with a three-peaked

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

CellPress

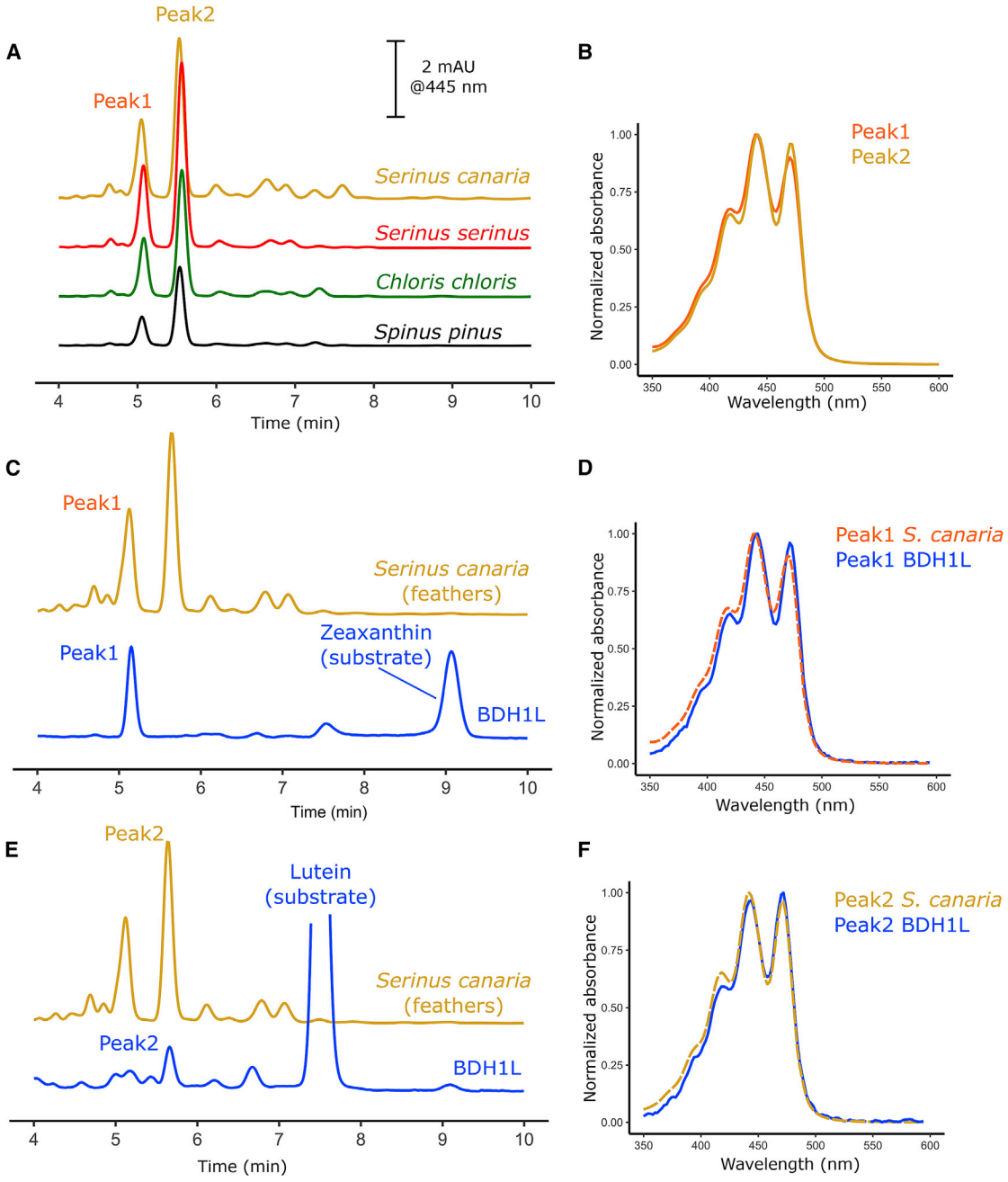


Figure 4. BDH1L alone converts dietary yellow carotenoids into ε,ε-carotenoids

(A) HPLC-DAD profiles of carotenoids extracted from yellow feathers of four species of true finches (Fringillidae).

(B) The absorbance spectra of peaks 1 and 2 from domestic canary (*Serinus canaria*).

(C and D) BDH1L converts zeaxanthin into an ε,ε-carotenoid with a retention time (C) and absorbance spectrum (D) closely matching those of the second most abundant carotenoid in canary feathers (peak 1).

(E and F) BDH1L converts lutein into an ε,ε-carotenoid with a retention time (E) and absorbance spectrum (F) closely matching those of the most abundant carotenoid in canary feathers (peak 2).

See also Figure S1.

absorption spectrum consistent with an ε,ε-carotenoid (Figures 3A and 3B). Peak 3 has the same retention time as the second most abundant carotenoid in the feathers of several true finch species (Fringillidae) (Figures 4A–4D). Mass spectrometry showed that peak 3 and the corresponding peak from canary feathers have identical molecular precursor ion and product ion fragmentation patterns (Figures S1E–S1H). The detection of an ion at *m/z* 597 suggests that this carotenoid

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

has the same molecular formula as astaxanthin ($C_{40}H_{52}O_4$), but the presence of two ε-rings implies that the functional groups likely occupy different positions than those in astaxanthin. Together, the HPLC retention time, the absorption spectrum, and the mass spectrometry data suggest that peak 3 is canary xanthophyll D or isoastaxanthin (i.e., 4,4′-dihydroxy-ε,ε-carotene-3,3′-dione).[10] This conclusion differs from that of a prior report which identified peak 1 (Figure 4A) as canary xanthophyll B (i.e., ε,ε-carotene-3,3′-dione).[16] We also found that expression of BDH1L alone results in conversion of lutein into an ε,ε-carotenoid with the same retention time and absorption spectrum as the most abundant carotenoid in the feathers of true finches (Figures 4E and 4F). The ability of BDH1L to convert zeaxanthin and lutein into ε,ε-carotenoids in transfected cells may explain the widespread co-occurrence of ε,ε-carotenoids and ketocarotenoids in the feathers of diverse bird species.[15] However, additional assays utilizing purified BDH1L protein will be required to determine whether this enzyme in isolation is sufficient for this conversion.

Upon closer examination of the products of BDH1L acting on zeaxanthin, we noted trace quantities of several additional carotenoids that were absent from the untransfected control cultures. Two of these products showed absorption spectra identical to those of α-doradexanthin and astaxanthin but with different retention times (data not shown). Thus, transfection of BDH1L alone into cultured mammalian cells appears to catalyze the production of minute quantities of ketocarotenoid from zeaxanthin. It is possible that endogenous cytochrome P450s in the cultured cells play a role in this conversion.

Next, we sought to determine the order in which CYP2J19 and BDH1L act on dietary yellow carotenoid substrates to produce ketocarotenoids. We first considered the possibility that BDH1L acts on the substrate prior to CYP2J19. To test this idea, we transfected cultured cells with CYP2J19 and provided peak 1 from canary feathers as substrate. Peak 1 from canary feather is identical to the product of BDH1L acting on zeaxanthin (Figures 4C and S1E–S1H) and therefore serves as an abundant and readily accessible source of this substrate. We saw no evidence of ketocarotenoid production by CYP2J19 (Figure S3A). We also provided CYP2J19-transfected cells with peak 2 from canary feathers as substrate. Again, we saw no evidence of ketocarotenoid production (Figure S3B). Next, we collected the product of CYP2J19 acting on zeaxanthin (peak 2 in Figures 3A and 3B) and provided it as substrate to BDH1L-transfected cells. This assay resulted in production of astaxanthin (Figures S3F and S3G). These findings suggest that CYP2J19 acts first on zeaxanthin to produce an intermediate which is then acted upon by BDH1L to generate astaxanthin. These results also suggest that ε,ε-carotenoids are not substrates for CYP2J19.

CYP2J19 was previously shown to be highly expressed in regenerating red feather follicles and in developing red cones.[21] Our transcriptome profiling data indicate that *BDH1L* is also enriched in red cones (Data S1). To determine whether *BDH1L* is expressed in regenerating red feathers, we performed *in situ* hybridization on sections of regenerating feather follicles from canaries exhibiting red and yellow carotenoid-based coloration. We found that *BDH1L* is expressed in an identical pattern to that of *CYP2J19* (Figure 5A), indicating that these genes are co-expressed in regenerating feather follicles. To determine whether these two genes are also involved in the regulation of red feather coloration, we reanalyzed published genetic data on the northern flicker (*Colaptes auratus*),[33] a species that exhibits two regional color morphs with yellow and red feathers (Figure 5B). The feathers of the yellow-shafted morph are pigmented with dietary yellow carotenoids, whereas the feathers of the red-shafted morph contain ketocarotenoids.[34] A genome-wide association study of yellow/red hybrid birds pinpointed multiple genomic regions associated with yellow/red variation,[33] including a locus on chromosome 8 which corresponds to *CYP2J19*. Our reanalysis confirmed the *CYP2J19* locus (Figures 5C and 5E) and found that a second locus (which the authors previously identified on chromosome 3) is centered on *BDH1L* (Figures 5D and 5F). We also revisited our own prior genetic analysis of the red factor canary,[21] where we first identified *CYP2J19* as a gene required for ketocarotenoid-based pigmentation. In that study, we detected a weak signal on chromosome 3 that also corresponds to the *BDH1L* locus. Thus, genetic studies implicate both *CYP2J19* and *BDH1L* in the control of ketocarotenoid pigmentation in birds.

Fish uses a similar mechanism of ketocarotenoid biosynthesis

Many fish species use yellow xanthophores and red erythrophores to pigment their skin.[35] Xanthophores frequently contain dietary yellow carotenoids, whereas erythrophores contain ketocarotenoids.[36] To better understand the molecular mechanisms of red coloration, we recently used RNA-seq to compare gene expression between fin regions containing only erythrophores or only xanthophores in a zebrafish relative, the pearl danio (*Danio albolineatus*).[36] We then screened 25 differentially expressed genes by knockout and identified two erythrophore-enriched genes that, when mutated, caused selective loss of ketocarotenoids: *cyp2ae2* and *bdh1a* (also known as *zgc:113142*) (Figure 6A). To evaluate the similarity between these two fish enzymes and avian CYP2J19 and BDH1L, we undertook phylogenetic analysis. We found that BDH1L and Bdh1a are both members of a well-defined short-chain dehydrogenase/reductase sub-family distinct from the BDH1 sub-family (Figure S4A), indicating that the fish and bird enzymes are orthologous. We also conducted phylogenetic analysis of cytochrome P450 enzymes related to CYP2J19 and Cyp2ae2. We found that CYP2J19 and Cyp2ae2 are closely related, although there are additional *Danio* P450s with a similar degree of relatedness to CYP2J19 (Figure S4B).

To determine whether Cyp2ae2 and Bdh1a can convert zeaxanthin into astaxanthin, we transfected these genes into cultured mammalian cells and provided zeaxanthin as substrate. Together, the enzymes produced astaxanthin (Figures 6B and 6C). Furthermore, when assayed alone, Cyp2ae2 generated a product (peak 2 in Figures 6B and 6C) with the same retention time and absorption spectrum as the product of CYP2J19 alone (peak 2 in Figures 3A and 3B), suggesting a similar mechanism of action. In contrast, we did not detect any product when Bdh1a was assayed alone (Figure 6B). To test whether these enzymes are sufficient to drive ketocarotenoid production *in vivo*, we expressed *D. albolineatus cyp2ae2* and *bdh1a* in zebrafish (*Danio rerio*), which lacks ketocarotenoid-based coloration. We

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

CellPress

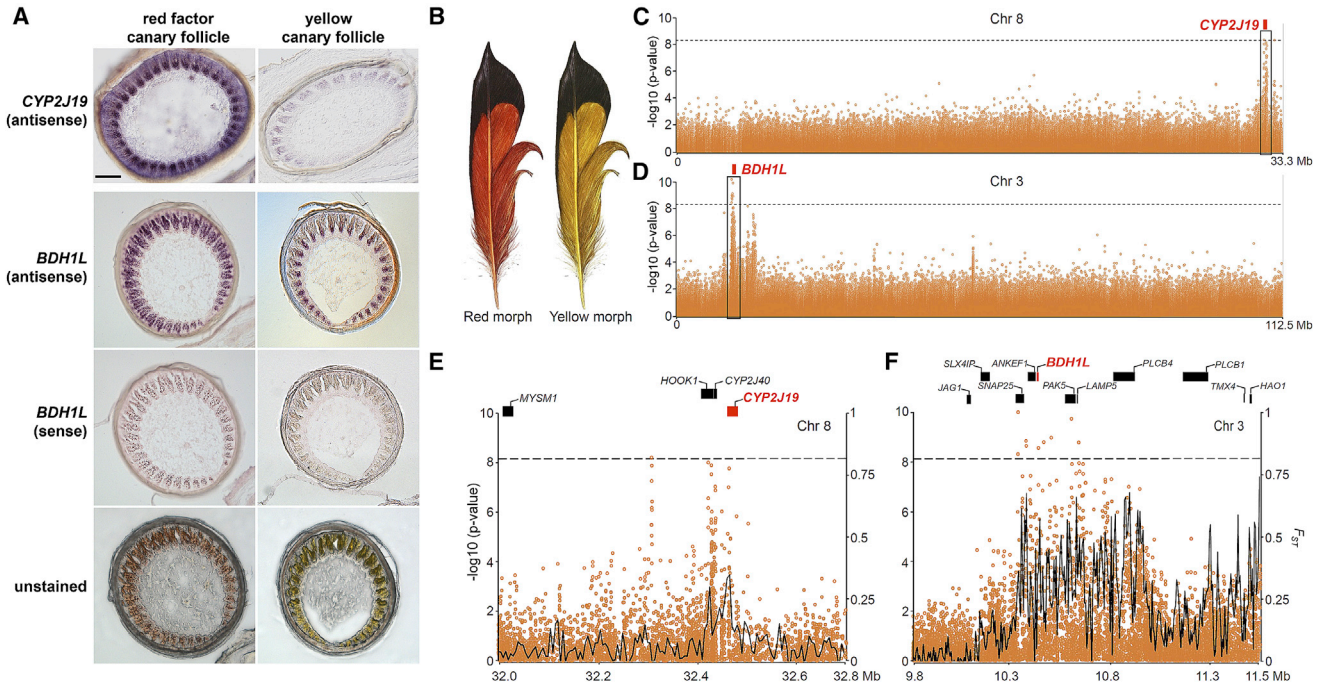


Figure 5. *CYP2J19* and *BDH1L* are co-expressed in regenerating feather follicles and genetically regulate red feather coloration

(A) *In situ* hybridization of regenerating feather follicles from red factor and yellow canaries (*Serinus canaria*). The *CYP2J19* images are from a prior publication[21] and are reproduced here for comparison. Scale bars, 100 μm.

(B–F) A prior genome-wide association study identified genomic regions associated with yellow/red feather color variation in the northern flicker (*Colaptes auratus*); data from 48 yellow/red hybrid individuals.[33]

(B) Schematic representation of tail feathers of red-shafted (left) and yellow-shafted (right) northern flickers.

(C–F) Candidate genomic regions encompassing *CYP2J19* and *BDH1L* are implicated in controlling carotenoid-based coloration differences in wing and tail feathers of northern flickers[33] (C and D, chromosome-level view; E and F, close-up of boxed regions in C and D, respectively). Each dot represents the $-\log_{10}$ transformation of Wald test p values for each individual SNP (the scale is shown on the left side of the y axis). The horizontal dashed line indicates the Bonferroni-corrected genome-wide significance ($p = 6.92 \times 10^{-9}$) based on the total number of tests (n = 7,229,704). In (E) and (F), the solid black line represents F_{ST} between allopatric populations of red-shafted and yellow-shafted flickers averaged in 25 kb windows, iterated in steps of 5 kb (the scale is shown on the right side of the y axis). Protein coding genes contained within each region are shown as black and red bars at the top of the graphs.

observed conversion of the normally yellow xanthophores into reddish orange erythrophore-like cells (Figures 6D–6G). HPLC-DAD analysis confirmed the presence of ketocarotenoids in the skin of the double transgenic fish, but not in the skin of fish expressing *cyp2ae2* or *bdh1a* individually (Figures 6H–6K). Taken together, these studies demonstrate that Cyp2ae2 and Bdh1a are both necessary and sufficient for ketocarotenoid biosynthesis and that *D. albolineatus* utilizes a two-enzyme mechanism very similar to that used by birds.

TTC39B enhances ketocarotenoid production

To identify additional factors involved in ketocarotenoid formation, we used whole-genome resequencing and genetic mapping to pinpoint causal mutations in "yellow" or "orange" morphs of three species of red-feathered birds. The different morphs in each species exhibited no discernible genetic structure with respect to color variation, a fact that facilitated genetic mapping (Figure S5). In two of these species, the diamond firetail (*Stagonopleura guttata*) and the star finch (*Bathilda ruficauda*), the mutations mapped to the *CYP2J19* locus and were not analyzed further (Figure S5). In a third species, the red-throated parrotfinch (*Erythrura psittacea*; Figures 7A and 7B), we found that all sequence variants significantly associated with red/orange coloration were located on chromosome Z scaffolds (Figure 7E). Controlled crosses confirmed that this polymorphism is sex linked and recessively inherited (data not shown). Next, we identified the minimal candidate interval where all orange individuals carried a common haplotype (Figure 7F). To pinpoint the causative gene in this interval, we used RNA-seq to compare gene expression in regenerating feather follicles of red and orange morphs. We found that none of the genes in this interval were differentially expressed between the two morphs (Table S2). However, using Illumina reads and Nanopore long-read sequencing technology, we detected an 18.1 kb duplication encompassing exons 3–11 of *TTC39B* that was present in homozygosity in orange, but not red, individuals (Figure 7G). To investigate the potential effects of this mutation on the splicing of *TTC39B*, we sequenced full-length transcripts from regenerating feather follicles and retina of both red and orange individuals using Nanopore technology. We found in both tissues a high proportion of abnormally spliced and mostly truncated transcripts in orange individuals (Figure S6). We therefore conclude that the duplication in *TTC39B* and the resultant reduction in the levels of functional transcripts likely explain the orange morph

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

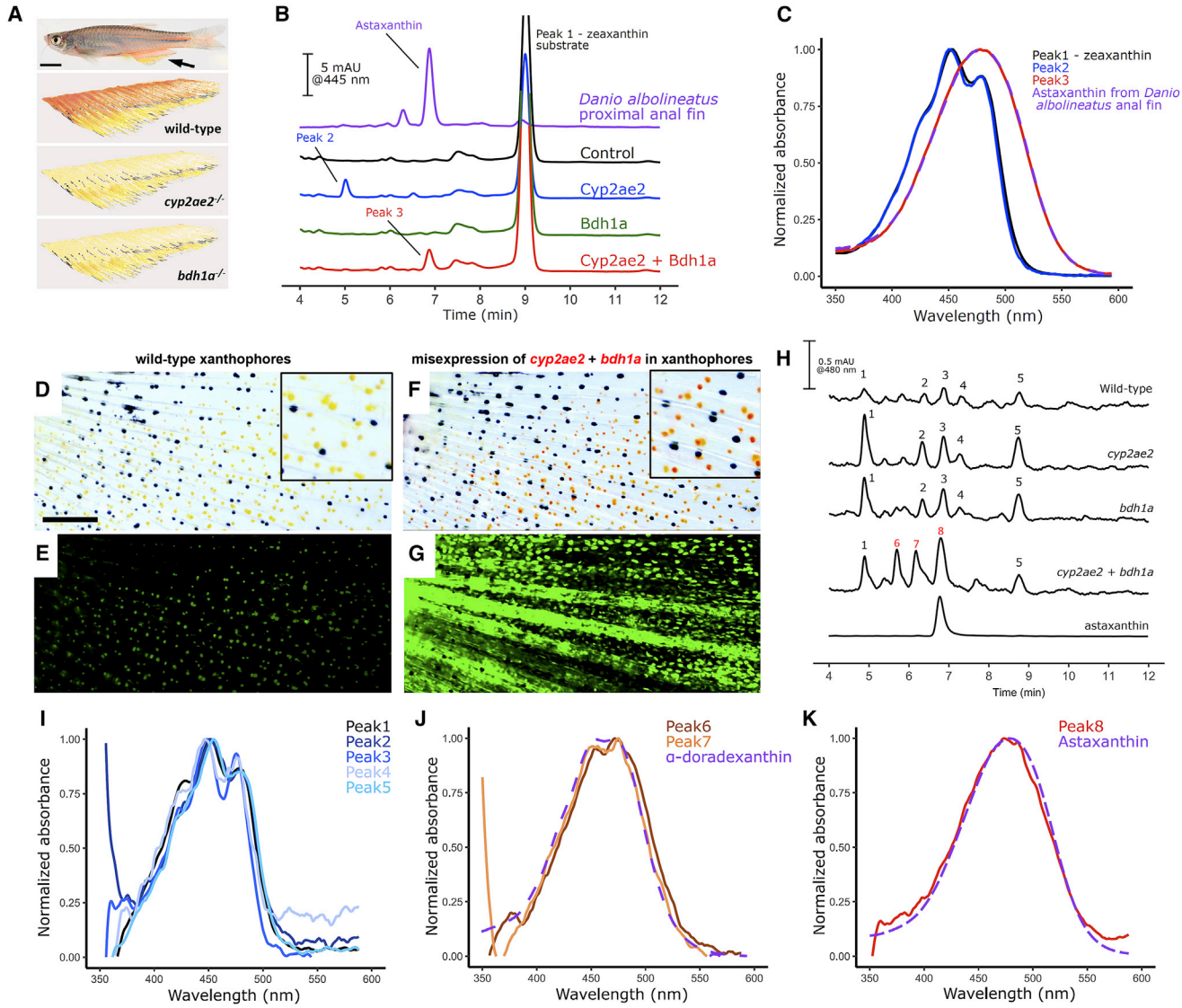


Figure 6. Fish use the same mechanism of ketocarotenoid biosynthesis

(A) The top panel shows an adult pearl danio (*Danio albolineatus*; arrow indicates anal fin). Scale bars, 5 mm. The bottom panels are schematic representations of anal fins of wild-type, *cyp2ae2*$^{-/-}$, and *bdh1a*$^{-/-}$ fish. The proximal half of wild-type fins are populated by orange ketocarotenoid-containing erythrophores that are severely reduced in *cyp2ae2*$^{-/-}$ and *bdh1a*$^{-/-}$ fish.[36]

(B and C) HPLC-DAD profiles (B) and absorption spectra (C) of carotenoids extracted from HEK293 cells transfected with pearl danio *cyp2ae2*, *bdh1a*, or both and provided with a zeaxanthin substrate.

(D) High-magnification view of the caudal fin of zebrafish (*Danio rerio*), which normally contain xanthophores (yellow) but lacks erythrophores. Black dots are melanophores. Size bar, 200 μm.

(E) Same field of view as in (D), showing xanthophore autofluorescence upon exposure to 488 nm light.

(F and G) Ectopic expression of pearl danio *cyp2ae2* and *bdh1a* confers orange coloration upon zebrafish xanthophores. Ectopic expression was achieved by co-expressing *cyp2ae2* and *bdh1a* under control of a heat-shock-inducible promoter along with nEOS, a green fluorescent protein. The fish carry the following transgenes: *hsp70l*::*cyp2ae2-2A-nEOS* and *hsp70l*::*bdh1a-2A-nEOS*.

(F) High-magnification view of the caudal fin of transgenic zebrafish after 2 weeks of heat shock.

(G) Same field of view as in (F), showing a marked increase in green fluorescence attributable to the expression of nEOS.

(H) HPLC-DAD profiles of caudal fin extracts from zebrafish ectopically co-expressing *cyp2ae2* and *bdh1a* contain novel peaks (6, 7, and 8) which are absent from wild-type fins. Numbers in black indicate peaks with absorption spectra similar to the corresponding numbered peaks in the wild-type trace.

(I) The absorption spectra of peaks 1–5 are typical of dietary yellow carotenoids. Only the spectra of peaks 1–5 from *cyp2ae2*-expressing transgenics are shown here.

(J) The absorption spectra of peaks 6 and 7 are similar to that of the ketocarotenoid α-doradexanthin.

(K) Peak 8 has the same retention time (see H) and absorption spectrum as an authentic astaxanthin standard.

See also Figure S4.

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

CellPress

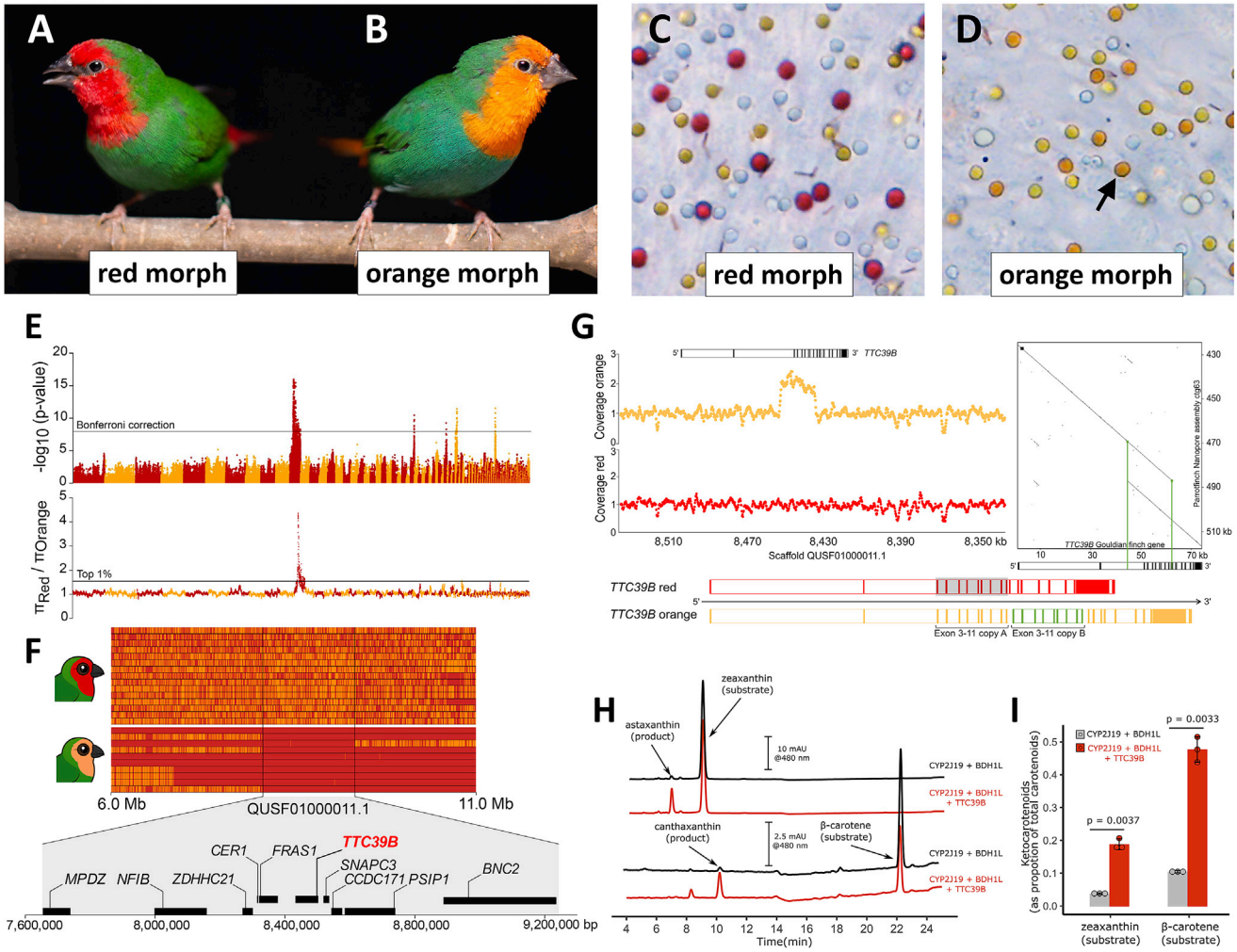


Figure 7. TTC39B enhances ketocarotenoid biosynthesis

(A and B) Images of wild-type (red morph) and mutant (orange morph) red-throated parrotfinches (*Erythrura psittacea*).

(C and D) Retinas of red and orange morphs viewed *en face*. The red oil droplets of the wild-type bird are replaced with orange droplets in the mutant.

(E and F) Genetic mapping of the *orange* mutation in red-throated parrotfinches.

(E) The upper panel summarizes genome-wide association analysis (red, n = 20; orange, n = 25). Each dot represents the $-\log_{10}$ transformation of likelihood ratio test (LRT) p values for each individual SNP. The horizontal line indicates the Bonferroni-corrected genome-wide significance ($p = 1.42 \times 10^{-8}$) based on the total number of tests (n = 3,514,701). The lower panel summarizes the ratio of nucleotide diversity (π) in red (wild type) and orange parrotfinches. Each dot represents the ratio in 50 kb windows, iterated in steps of 10 kb. Scaffolds smaller than 500 kb are omitted for clarity.

(F) Individual genotypes across the candidate region. Each line corresponds to one male individual and each column to a variant. Red and orange indicate homozygous and heterozygous positions, respectively. The vertical black lines flank a region of high homozygosity in orange parrotfinches. The protein coding genes contained within the minimum common haplotype shared among orange individuals are shown at the bottom. Parrotfinch icons are courtesy of Birdorable.com.

(G) The top left panel summarizes the normalized number of mapped reads in non-overlapping windows of 1 kb averaged across all red (bottom plot) and orange (top plot) males. Read depth is approximately doubled throughout a stretch of ~18 kb overlapping exons 3–11 of *TTC39B*, consistent with a duplication (QUSF01000011.1: 8,433,112–8,451,216 bp). The right panel represents a dot plot of the alignment between the Gouldian finch *TTC39B* genome sequence and the homologous scaffold from a draft genome assembly of an orange parrotfinch individual obtained using Nanopore reads. Part of the sequence of the orange individual aligns twice to the same segment in the Gouldian finch genome. The bottom panel shows a schematic representation of the wild-type *TTC39B* gene model and the orange gene model with the duplicated region highlighted by a gray box.

(H) HPLC elution profiles of carotenoids extracted from HEK293 cells transfected with the indicated constructs and provided with zeaxanthin or β-carotene as substrate.

(I) Quantitation of ketocarotenoid production by co-expression of TTC39B in cultured cells. Bars represent the mean ± SD (n = 3) proportion of ketocarotenoid product relative to total carotenoid content (i.e., substrate + product) in the transfected cells; points represent the individual replicates. Conditions were compared with Welch's two sample t test.

See also Table S2 and Figures S5–S7.

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

phenotype. Of note, this gene has previously been implicated in carotenoid-based coloration in other bird and fish species.[37,38]

The red-throated parrotfinch normally has bright red feathers covering the head and throat (Figure 7A). In the orange morph, these feathers are yellowish orange (Figure 7B). Similarly, the red cone photoreceptors of the wild-type bird contain bright red oil droplets, whereas these oil droplets are orange in the mutant (Figures 7C and 7D). We used HPLC-DAD analysis to show that ketocarotenoid levels are markedly reduced in the feathers and retinas of the orange morph compared with wild type (Figure S7). Zeaxanthin levels are also modestly reduced in the retina, but not to the same extent as astaxanthin levels. In contrast, plasma and liver carotenoid levels are unchanged in the mutant (Figure S7). High-density lipoprotein (HDL) receptor is the main plasma transporter of xanthophyll carotenoids in vertebrates,[39–41] and a genome-wide association study in humans previously demonstrated a role for *TTC39B* in control of plasma HDL levels.[42] We therefore measured HDL, LDL, and triglyceride levels in plasma and liver of wild-type and mutant birds. We found no differences between the two (Figure S7). We also found no differences in the ratios of saturated, mono-unsaturated, and poly-unsaturated fatty acids in body fat of wild type and mutant (Figure S7). We infer from these findings that the *TTC39B* mutation is not likely to exert its effect via changes in HDL or total circulating carotenoid levels. Instead, we hypothesize that *TTC39B* might be acting at the sites of ketocarotenoid biosynthesis (i.e., in the feather follicles and the red cone). Consistent with this idea, *TTC39B* is expressed at ~10-fold higher levels in the chicken red cone than that in the other cone subtypes at P15 (Figure 2H; Data S1).

To further evaluate the function of TTC39B, we co-transfected the chicken gene along with CYP2J19 and BDH1L into cultured mammalian cells and provided the cells with either zeaxanthin or β-carotene as substrate. We found that TTC39B potently enhanced the conversion of zeaxanthin and β-carotene into the corresponding ketocarotenoids (Figures 7H and 7I). We propose that TTC39B may act by facilitating the transport of carotenoid substrates to or from the site of enzymatic conversion.

DISCUSSION

Red ketocarotenoid pigments are deposited in the integument and retina of many vertebrates and are critical for coloration and color vision. In the present study, we describe a mechanism of ketocarotenoid biosynthesis in vertebrates that requires CYP2J19 and BDH1L and is enhanced by TTC39B. We find that both birds and fish employ a similar mechanism, suggesting that this enzymatic pathway may be widely utilized by vertebrates. The observation that an analogous mechanism for carotenoid conversion is present in such highly divergent taxa, where it presumably evolved independently, suggests that this biochemical pathway is particularly "accessible" in evolution. Indeed, repeated gains and losses of carotenoid-based coloration are frequent in birds,[15,43,44] and there have been multiple independent origins of ketocarotenoid-based feather pigmentation in diverse clades.[45,46] We showed that chicken CYP2J19 and BDH1L can convert diverse dietary yellow carotenoids into ketocarotenoids, although this species normally only produces astaxanthin in the retina. Thus, these enzymes appear to be pre-adapted for deployment in the integument, where substrate promiscuity would enable utilization of diverse dietary carotenoids for ketocarotenoid production. Moreover, genetic variation around *CYP2J19*, *BDH1L*, and *TTC39B*, or some combination of these genes, is associated with the control of red ketocarotenoid production in nearly every wild and domesticated bird species studied to date.[21,33,37] We propose that the molecular simplicity of this mechanism explains why transitions from yellow-to-red carotenoid pigmentation, and vice versa, occur so frequently in vertebrates.

Ketocarotenoids and ε,ε-carotenoids co-occur in the feathers of diverse species of birds,[15,16] but the reason for their co-occurrence has long been a mystery. The discovery that BDH1L plays a role in converting dietary carotenoids into ε,ε-carotenoids offers a potential explanation. The evolution of the ability to produce ketocarotenoids in feather follicles appears to require CYP2J19 and BDH1L co-expression, perhaps via acquisition of follicle-specific enhancers at these loci. Subsequent loss of follicular CYP2J19 expression (with maintenance of BDH1L expression) in individual species would then lead to the exclusive production of ε,ε-carotenoids. A recent study in true finches provides support for this idea.[15] The authors suggest that the capacity to synthesize ketocarotenoids may have evolved only once in this family. Multiple descendant species of the primordial "red" bird later reverted back to yellow feather coloration, but in all cases, those species were found to primarily deposit modified ε,ε-carotenoids in their feathers, not dietary yellow carotenoids. We therefore propose that evolutionary loss of follicular CYP2J19 expression (with sustained BDH1L expression) in specific lineages might explain the present-day distribution of ε,ε-carotenoids in true finches and perhaps other families.

Carotenoid coloration plays a key role in social interactions including mate choice,[1,46–48] and some authors have proposed that ketocarotenoid production might be linked to underlying metabolic processes and therefore represent an honest signal of individual quality.[49–51] The discovery of an enzymatic pathway for ketocarotenoid biosynthesis has important implications for understanding the potential connections between coloration and basic metabolic pathways. Furthermore, our findings provide a basis for interpreting the results of genetic mapping studies aimed at identifying loci that regulate color patterning. More generally, our work opens new avenues for understanding the mechanisms underlying the evolution of visual ornaments and color vision.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell line
  - Animals

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.08.013.

ACKNOWLEDGMENTS

We thank members of the Corbo laboratory for feedback and discussion. We also thank Jocelyn Hudon, Charles Brown, Bruno Araújo, Miguel Monteiro, and Vítor Lopes for feather or blood samples. We thank Meadow Hansen, Desirae Gonzales, Brooke Joski, and Hannah Reeb for laboratory assistance, and we also thank Erica Lantelme and the Flow Cytometry Core in the Department of Pathology and Immunology at Washington University in St. Louis for FACS method development and services. The Genome Technology Access Core (GTAC) in the Department of Genetics at Washington University in St. Louis provided next-generation sequencing services. This work was supported by R01 EY030075 (to J.C.C.), NIH R35 GM122471 (to D.M.P.), NSF IOS 2037739 (to M.B.T.), as well as lab startup funds, Faculty Development Summer Fellowships, and the Undergraduate Research Challenge program at the University of Tulsa (to M.B.T.). Part of the carotenoid sample analysis was supported by USDA NIFA Hatch #W4122 (to R.E. Kopec) and NIH grant P30 CA016058 (to the Ohio State University and OSUCCC). This work was also supported by research contracts from the Fundação para a Ciência e Tecnologia (FCT) CEECINST/00014/2018/CP1512/CT0002 (to M.C.), 2020.01405.CEECIND/CP1601/CT0011 (to P.A.), and 2020.01494.CEECIND (to P.M.A.); by the European Research Council (ERC; to M.C.) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 101000504); by Portuguese national funds Transitory Norm contract (DL57/2016/CP1440/CT0006; to R.J.L.); and by research fellowships PD/BD/114042/2015 (to M.A.G.), SFRH/BD/147030/2019 (to C.I.M.), and PD/BD/128492/2017 (to P.P.) under the auspices of the Biodiversity, Genetics, and Evolution (BIODIV) PhD program.

AUTHOR CONTRIBUTIONS

M.B.T., M.C., and J.C.C. conceived the project. M.B.T., D.J.S., M.E.H., M.C., C.I.M., P.M.A., P.P., S.A., P.A., M.A.G., R.J.L., I.V., S.Z., R.E. Kopec, R.E. Koch, D.H., D.M.P., Y.L., G.D.S., C.A.M., Y.O., D.M., and J.C.C. performed the experiments and/or analyzed the data. M.B.T., M.C., and J.C.C. supervised and administered the project. M.B.T., M.C., R.E. Kopec, D.M.P., and J.C.C. acquired funding for the project. M.B.T., M.C., and J.C.C. wrote the manuscript with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 27, 2022
Revised: June 19, 2022
Accepted: August 8, 2022
Published: August 31, 2022

REFERENCES

1. Darwin, C. (1871). The Descent of Man, and Selection in Relation to Sex (John Murray).
2. Cuthill, I.C., Allen, W.L., Arbuckle, K., Caspers, B., Chaplin, G., Hauber, M.E., Hill, G.E., Jablonski, N.G., Jiggins, C.D., Kelber, A., et al. (2017). The biology of color. Science *357*. eaan0221. https://doi.org/10.1126/science.aan0221.
3. Svensson, P.A., and Wong, B.B.M. (2011). Carotenoid-based signals in behavioural ecology: a review. Behaviour *148*, 131–189. https://doi.org/10.1163/000579510X548673.
4. Hill, G.E., and McGraw, K.J. (2006). Bird coloration. In Mechanisms and Measurements, *1* (Harvard University Press).
5. Maoka, T. (2020). Carotenoids as natural functional pigments. J. Nat. Med. *74*, 1–16. https://doi.org/10.1007/s11418-019-01364-x.
6. Sefc, K.M., Brown, A.C., and Clotfelter, E.D. (2014). Carotenoid-based coloration in cichlid fishes. Comp. Biochem. Physiol. A Mol. Integr. Physiol. *173C*, 42–51. https://doi.org/10.1016/j.cbpa.2014.03.006.
7. Goldsmith, T.H., Collins, J.S., and Licht, S. (1984). The cone oil droplets of avian retinas. Vision Res *24*, 1661–1671.
8. Toomey, M.B., and Corbo, J.C. (2017). Evolution, development and function of vertebrate cone oil droplets. Front. Neural Circuits *11*, 97. https://doi.org/10.3389/fncir.2017.00097.
9. Vorobyev, M. (2003). Coloured oil droplets enhance colour discrimination. Proc. Biol. Sci. *270*, 1255–1261.
10. LaFountain, A.M., Prum, R.O., and Frank, H.A. (2015). Diversity, physiology, and evolution of avian plumage carotenoids and the role of carotenoid–protein interactions in plumage color appearance. Arch. Biochem. Biophys. *572*, 201–212. https://doi.org/10.1016/j.abb.2015.01.016.
11. Davies, B.H. (1985). Carotenoid metabolism in animals: a biochemist's view. Pure Appl. Chem. *57*, 679–684. https://doi.org/10.1351/pac198557050679.
12. Amengual, J., Widjaja-Adhi, M.A.K., Rodriguez-Santiago, S., Hessel, S., Golczak, M., Palczewski, K., and von Lintig, J. (2013). Two carotenoid oxygenases contribute to mammalian provitamin A metabolism. J. Biol. Chem. *288*, 34081–34096. https://doi.org/10.1074/jbc.M113.501049.
13. Toomey, M.B., Lind, O., Frederiksen, R., Curley, R.W., Jr., Riedl, K.M., Wilby, D., Schwartz, S.J., Witt, C.C., Harrison, E.H., Roberts, N.W., et al. (2016). Complementary shifts in photoreceptor spectral tuning unlock the full adaptive potential of ultraviolet vision in birds. eLife *5*, e15675, https://doi.org/10.7554/eLife.15675.

14. Stradi, R., Celentano, G., Boles, M., and Mercato, F. (1997). Carotenoids in bird plumage: the pattern in a series of red-pigmented Carduelinae. Comp. Biochem. Mol. Biol. *117*, 85–91. https://doi.org/10.1016/S0305-0491(96)00271-4.

15. Ligon, R.A., Simpson, R.K., Mason, N.A., Hill, G.E., and McGraw, K.J. (2016). Evolutionary innovation and diversification of carotenoid-based pigmentation in finches. Evolution *70*, 2839–2852. https://doi.org/10.1111/evo.13093.

16. Stradi, R., Celentano, G., Rossi, E., Rovati, G., and Pastore, M. (1995). Carotenoids in bird plumage I. The carotenoid pattern in a series of palearctic Carduelinae. Comp. Biochem. Physiol. B *110*, 131–143.

17. Mays, H.L., Jr., McGraw, K.J., Ritchison, G., Cooper, S., Rush, V., and Parker, R.S. (2004). Sexual dichromatism in the yellow-breasted chat *Icteria virens*: spectrophotometric analysis and biochemical basis. J. Avian Biol. *35*, 125–134. https://doi.org/10.1111/j.0908-8857.2004.03101.x.

18. McGraw, K.J., Hill, G.E., Stradi, R., and Parker, R.S. (2001). The influence of carotenoid acquisition and utilization on the maintenance of species-typical plumage pigmentation in male American goldfinches (*Carduelis tristis*) and northern cardinals (*Cardinalis cardinalis*). Physiol. Biochem. Zool. *74*, 843–852. https://doi.org/10.1086/323797.

19. Toews, D.P.L., Hofmeister, N.R., and Taylor, S.A. (2017). The evolution and genetics of carotenoid processing in animals. Trends Genet *33*, 171–182. https://doi.org/10.1016/j.tig.2017.01.002.

20. Price-Waldman, R., and Stoddard, M.C. (2021). Avian coloration genetics: recent advances and emerging questions. J. Hered. *112*, 395–416. https://doi.org/10.1093/jhered/esab015.

21. Lopes, R.J., Johnson, J.D., Toomey, M.B., Ferreira, M.S., Araujo, P.M., Melo-Ferreira, J., Andersson, L., Hill, G.E., Corbo, J.C., and Carneiro, M. (2016). Genetic basis for red coloration in birds. Curr. Biol. *26*, 1427–1434. https://doi.org/10.1016/j.cub.2016.03.076.

22. Mundy, N.I., Stapley, J., Bennison, C., Tucker, R., Twyman, H., Kim, K.W., Burke, T., Birkhead, T.R., Andersson, S., and Slate, J. (2016). Red carotenoid coloration in the zebra finch is controlled by a cytochrome P450 gene Cluster. Curr. Biol. *26*, 1435–1440. https://doi.org/10.1016/j.cub.2016.04.047.

23. Toomey, M.B., Collins, A.M., Frederiksen, R., Cornwall, M.C., Timlin, J.A., and Corbo, J.C. (2015). A complex carotenoid palette tunes avian colour vision. J. R. Soc. Interface *12*, 20150563, https://doi.org/10.1098/rsif.2015.0563.

24. Bhosale, P., Serban, B., Zhao, D.Y., and Bernstein, P.S. (2007). Identification and metabolic transformations of carotenoids in ocular tissues of the Japanese quail Coturnix japonica. Biochemistry *46*, 9050–9057. https://doi.org/10.1021/bi700558f.

25. Kram, Y.A., Mantey, S., and Corbo, J.C. (2010). Avian cone photoreceptors tile the retina as five independent, self-organizing mosaics. PLoS One *5*, e8992, https://doi.org/10.1371/journal.pone.0008992.

26. Enright, J.M., Lawrence, K.A., Hadzic, T., and Corbo, J.C. (2015). Transcriptome profiling of developing photoreceptor subtypes reveals candidate genes involved in avian photoreceptor diversification. J. Comp. Neurol. *523*, 649–668. https://doi.org/10.1002/cne.23702.

27. Pimentel, H., Bray, N.L., Puente, S., Melsted, P., and Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. Nat. Methods *14*, 687–690. https://doi.org/10.1038/nmeth.4324.

28. Kaufman, L., and Rousseeuw, P.J. (2009). Chapter 6. Divisive analysis (program DIANA). In Finding Groups in Data: An Introduction to Cluster Analysis (Wiley), pp. 253–279.

29. Kavanagh, K.L., Jörnvall, H., Persson, B., and Oppermann, U. (2008). Medium- and short-chain dehydrogenase/reductase gene and protein families: the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. Cell. Mol. Life Sci. *65*, 3895–3906. https://doi.org/10.1007/s00018-008-8588-y.

30. Parker, R.O., and Crouch, R.K. (2010). Retinol dehydrogenases (RDHs) in the visual cycle. Exp. Eye Res. *91*, 788–792. https://doi.org/10.1016/j.exer.2010.08.013.

31. Puchalska, P., and Crawford, P.A. (2017). Multi-dimensional roles of ketone bodies in fuel metabolism, signaling, and therapeutics. Cell Metab *25*, 262–284. https://doi.org/10.1016/j.cmet.2016.12.022.

32. Schiedt, K. (1990). New aspects of carotenoid metabolism in animals. In Carotenoids: Chemistry and Biology, N.I. Krinsky, ed. (Plenum Press).

33. Aguillon, S.M., Walsh, J., and Lovette, I.J. (2021). Extensive hybridization reveals multiple coloration genes underlying a complex plumage phenotype. Proc. Biol. Sci. *288*, 20201805, https://doi.org/10.1098/rspb.2020.1805.

34. Hudon, J., Wiebe, K.L., Pini, E., and Stradi, R. (2015). Plumage pigment differences underlying the yellow-red differentiation in the Northern Flicker (*Colaptes auratus*). Comp. Biochem. Physiol. B Biochem. Mol. Biol. *183*, 1–10. https://doi.org/10.1016/j.cbpb.2014.12.006.

35. Parichy, D.M. (2021). Evolution of pigment cells and patterns: recent insights from teleost fishes. Curr. Opin. Genet. Dev. *69*, 88–96. https://doi.org/10.1016/j.gde.2021.02.006.

36. Huang, D., Lewis, V.M., Foster, T.N., Toomey, M.B., Corbo, J.C., and Parichy, D.M. (2021). Development and genetics of red coloration in the zebrafish relative *Danio albolineatus*. eLife *10*, e70253, https://doi.org/10.7554/eLife.70253.

37. Hooper, D.M., Griffith, S.C., and Price, T.D. (2019). Sex chromosome inversions enforce reproductive isolation across an avian hybrid zone. Mol. Ecol. *28*, 1246–1262. https://doi.org/10.1111/mec.14874.

38. Ahi, E.P., Lecaudey, L.A., Ziegelbecker, A., Steiner, O., Goessler, W., and Sefc, K.M. (2020). Expression levels of the tetratricopeptide repeat protein gene ttc39b covary with carotenoid-based skin colour in cichlid fish. Biol. Lett. *16*, 20200629, https://doi.org/10.1098/rsbl.2020.0629.

39. Harrison, E.H. (2019). Mechanisms of transport and delivery of vitamin A and carotenoids to the retinal pigment epithelium. Mol. Nutr. Food Res. *63*, e1801046, https://doi.org/10.1002/mnfr.201801046.

40. Connor, W.E., Duell, P.B., Kean, R., and Wang, Y. (2007). The prime role of HDL to transport lutein into the retina: evidence from HDL-deficient WHAM chicks having a mutant ABCA1 transporter. Invest. Ophthalmol. Vis. Sci. *48*, 4226–4231. https://doi.org/10.1167/iovs.06-1275.

41. Toomey, M.B., Lopes, R.J., Araújo, P.M., Johnson, J.D., Gazda, M.A., Afonso, S., Mota, P.G., Koch, R.E., Hill, G.E., Corbo, J.C., and Carneiro, M. (2017). High-density lipoprotein receptor SCARB1 is required for carotenoid coloration in birds. Proc. Natl. Acad. Sci. USA *114*, 5219–5224. https://doi.org/10.1073/pnas.1700751114.

42. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature *466*, 707–713. https://doi.org/10.1038/nature09270.

43. Thomas, D.B., McGraw, K.J., Butler, M.W., Carrano, M.T., Madden, O., and James, H.F. (2014). Ancient origins and multiple appearances of carotenoid-pigmented feathers in birds. Proc. Biol. Sci. *281*, 20140806, https://doi.org/10.1098/rspb.2014.0806.

44. Prum, R.O., LaFountain, A.M., Berro, J., Stoddard, M.C., and Frank, H.A. (2012). Molecular diversity, metabolic transformation, and evolution of carotenoid feather pigments in cotingas (Aves: Cotingidae). J. Comp. Physiol. B *182*, 1095–1116. https://doi.org/10.1007/s00360-012-0677-4.

45. Friedman, N.R., McGraw, K.J., and Omland, K.E. (2014). Evolution of carotenoid pigmentation in caciques and meadowlarks (Icteridae): repeated gains of red plumage coloration by carotenoid C4-oxygenation. Evolution *68*, 791–801. https://doi.org/10.1111/evo.12304.

46. Hill, G.E., and McGraw, K.J. (2006). Bird coloration. In Function and Evolution, *2* (Harvard University Press).

47. Hamilton, D.G., Whiting, M.J., and Pryke, S.R. (2013). Fiery frills: carotenoid-based coloration predicts contest success in frillneck lizards. Behav. Ecol. *24*, 1138–1149. https://doi.org/10.1093/beheco/art041.

48. Pryke, S.R., and Griffith, S.C. (2006). Red dominates black: agonistic signalling among head morphs in the colour polymorphic Gouldian finch. Proc. Biol. Sci. *273*, 949–957. https://doi.org/10.1098/rspb.2005.3362.

49. Weaver, R.J., Santos, E.S.A., Tucker, A.M., Wilson, A.E., and Hill, G.E. (2018). Carotenoid metabolism strengthens the link between feather coloration and individual quality. Nat. Commun. *9*, 73. https://doi.org/10.1038/s41467-017-02649-z.

50. Weaver, R.J., Koch, R.E., and Hill, G.E. (2017). What maintains signal honesty in animal colour displays used in mate choice? Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, https://doi.org/10.1098/rstb.2016.0343.

51. Powers, M.J., and Hill, G.E. (2021). A review and assessment of the shared-pathway hypothesis for the maintenance of signal honesty in red ketocarotenoid-based coloration. Integr. Comp. Biol. *61*, 1811–1826. https://doi.org/10.1093/icb/icab056.

52. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527. https://doi.org/10.1038/nbt.3519.

53. Melsted, P., Ntranos, V., and Pachter, L. (2019). The barcode, UMI, set format and BUStools. Bioinformatics Oxf. Engl. *35*, 4472–4473. https://doi.org/10.1093/bioinformatics/btz279.

54. Melsted, P., Booeshaghi, A.S., Liu, L., Gao, F., Lu, L., Min, K.H.J., da Veiga Beltrame, E., Hjörleifsson, K.E., Gehring, J., and Pachter, L. (2021). Modular, efficient and constant-memory single-cell RNA-seq preprocessing. Nat. Biotechnol. *39*, 813–818. https://doi.org/10.1038/s41587-021-00870-2.

55. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics Oxf. Engl. *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

56. Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. *37*, 907–915. https://doi.org/10.1038/s41587-019-0201-4.

57. Liao, Y., Smyth, G.K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res *47*, e47. https://doi.org/10.1093/nar/gkz114.

58. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

59. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res *43*, e47. https://doi.org/10.1093/nar/gkv007.

60. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

61. Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics Oxf. Engl. *30*, 3276–3278. https://doi.org/10.1093/bioinformatics/btu531.

62. Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics Oxf. Engl. *35*, 4453–4455. https://doi.org/10.1093/bioinformatics/btz305.

63. Darriba, D., Posada, D., Kozlov, A.M., Stamatakis, A., Morel, B., and Flouri, T. (2020). ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. Mol. Biol. Evol. *37*, 291–294. https://doi.org/10.1093/molbev/msz189.

64. Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics Oxf. Engl. *25*, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

65. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics Oxf. Engl. *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

66. Meisner, J., and Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. Genetics *210*, 719–731. https://doi.org/10.1534/genetics.118.301336.

67. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. BMC Bioinformatics *15*, 356. https://doi.org/10.1186/s12859-014-0356-4.

68. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics Oxf. Engl. *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

69. Garrison, E., Kronenberg, Z.N., Dawson, E.T., Pedersen, B.S., and Prins, P. (2021). Vcflib and tools for processing the VCF variant call format. Preprint at bioRxiv. https://doi.org/10.1101/2021.05.21.445151.

70. Browning, B.L., Zhou, Y., and Browning, S.R. (2018). A one-penny imputed genome from next-generation reference panels. Am. J. Hum. Genet. *103*, 338–348. https://doi.org/10.1016/j.ajhg.2018.07.015.

71. Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. *44*, 821–824. https://doi.org/10.1038/ng.2310.

72. Kinsella, R.J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., et al. (2011). Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database *2011*. bar030. https://doi.org/10.1093/database/bar030.

73. Layer, R.M., Chiang, C., Quinlan, A.R., and Hall, I.M. (2014). LUMPY: a probabilistic framework for structural variant discovery. Genome Biol *15*, R84. https://doi.org/10.1186/gb-2014-15-6-r84.

74. Sedlazeck, F.J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A., and Schatz, M.C. (2018). Accurate detection of complex structural variations using single-molecule sequencing. Nat. Methods *15*, 461–468. https://doi.org/10.1038/s41592-018-0001-7.

75. Noé, L., and Kucherov, G. (2005). YASS: enhancing the sensitivity of DNA similarity search. Nucleic Acids Res *33*, W540–W543. https://doi.org/10.1093/nar/gki478.

76. Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics Oxf. Engl. *34*, 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

77. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). Blast+: architecture and applications. BMC Bioinformatics *10*, 421. https://doi.org/10.1186/1471-2105-10-421.

78. National Research Council (2011). Guide for the Care and Use of Laboratory Animals Eighth Edition (The National Academies Press).

79. Murphy, D.P., Hughes, A.E., Lawrence, K.A., Myers, C.A., and Corbo, J.C. (2019). Cis-regulatory basis of sister cell type divergence in the vertebrate retina. eLife *8*, e48216, https://doi.org/10.7554/eLife.48216.

80. Hou, Z., Jiang, P., Swanson, S.A., Elwell, A.L., Nguyen, B.K., Bolin, J.M., Stewart, R., and Thomson, J.A. (2015). A cost-effective RNA sequencing protocol for large-scale gene expression studies. Sci. Rep. *5*, 9570. https://doi.org/10.1038/srep09570.

81. Trimarchi, J.M., Stadler, M.B., Roska, B., Billings, N., Sun, B., Bartch, B., and Cepko, C.L. (2007). Molecular heterogeneity of developing retinal ganglion and amacrine cells revealed through single cell gene expression profiling. J. Comp. Neurol. *502*, 1047–1065.

82. Prado-Cabrero, A., Saefurahman, G., and Nolan, J.M. (2020). Stereochemistry of astaxanthin biosynthesis in the marine harpacticoid copepod Tigriopus californicus. Mar. Drugs *18*, 506. https://doi.org/10.3390/md18100506.

83. Moretti, V.M., Mentasti, T., Bellagamba, F., Luzzana, U., Caprino, F., Turchini, G.M., Giani, I., and Valfrè, F. (2006). Determination of astaxanthin stereoisomers and colour attributes in flesh of rainbow trout (*Oncorhynchus mykiss*) as a tool to distinguish the dietary pigmentation source. Food Addit. Contam. *23*, 1056–1063. https://doi.org/10.1080/02652030600838399.

84. Kwan, K.M., Fujimoto, E., Grabher, C., Mangum, B.D., Hardy, M.E., Campbell, D.S., Parant, J.M., Yost, H.J., Kanki, J.P., and Chien, C.B.

(2007). The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. Dev. Dyn. *236*, 3088–3099. https://doi.org/10.1002/dvdy.21343.

85. Tan, J.A., and Mikheyev, A.S. (2016). A scaled-down workflow for Illumina shotgun sequencing library preparation: lower input and improved performance at small fraction of the cost. PeerJ *4*, v2471.
86. Toomey, M.B., Marques, C.I., Andrade, P., Araújo, P.M., Sabatino, S., Gazda, M.A., Afonso, S., Lopes, R.J., Corbo, J.C., and Carneiro, M. (2018). A non-coding region near follistatin controls head colour polymorphism in the Gouldian finch. Proc. Biol. Sci. *285*, 20181788, https://doi.org/10.1098/rspb.2018.1788.
87. Warren, W.C., Clayton, D.F., Ellegren, H., Arnold, A.P., Hillier, L.W., Künstner, A., Searle, S., White, S., Vilella, A.J., Fairley, S., et al. (2010). The genome of a songbird. Nature *464*, 757–762. https://doi.org/10.1038/nature08819.
88. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics Oxf. Engl. *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
89. Enbody, E.D., Sprehn, C.G., Abzhanov, A., Bi, H., Dobreva, M.P., Osborne, O.G., Rubin, C.J., Grant, P.R., Grant, B.R., and Andersson, L. (2021). A multispecies BCO2 beak color polymorphism in the Darwin's finch radiation. Curr. Biol. *31*, 5597–5604.e7. https://doi.org/10.1016/j.cub.2021.09.085.
90. Viegas, I., Rito, J., Jarak, I., Leston, S., Caballero-Solares, A., Metón, I., Pardal, M.A., Baanante, I.V., and Jones, J.G. (2015). Contribution of dietary starch to hepatic and systemic carbohydrate fluxes in European seabass (Dicentrarchus labrax L.). Br. J. Nutr. *113*, 1345–1354. https://doi.org/10.1017/S0007114515000574.
91. Matyash, V., Liebisch, G., Kurzchalia, T.V., Shevchenko, A., and Schwudke, D. (2008). Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. J. Lipid Res. *49*, 1137–1146. https://doi.org/10.1194/jlr.D700041-JLR200.
92. Duarte, J.A., Carvalho, F., Pearson, M., Horton, J.D., Browning, J.D., Jones, J.G., and Burgess, S.C. (2014). A high-fat diet suppresses de novo lipogenesis and desaturation but not elongation and triglyceride synthesis in mice. J. Lipid Res. *55*, 2541–2553. https://doi.org/10.1194/jlr.M052308.
93. R Core Team. (2021). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).
94. Yi, L., Pimentel, H., Bray, N.L., and Pachter, L. (2018). Gene-level differential analysis at transcript-level resolution. Genome Biol *19*, 53. https://doi.org/10.1186/s13059-018-1419-z.
95. Cingolani, P., Platts, A., Wang, le L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly *6*, 80–92. https://doi.org/10.4161/fly.19695.
96. Ruan, J., and Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. Nat. Methods *17*, 155–158. https://doi.org/10.1038/s41592-019-0669-3.

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

CellPress

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial strains		
E. coli – NEB5α	New England Biolabs	C2987H
Biological samples		
Danio albolineatus – fin tissue	David Parichy Laboratory, University of Virginia	N/A
Serinus serinus - feathers	Licensed breeders in Portugal	N/A
Chloris chloris - feathers	Licensed breeders in Portugal	N/A
Spinus pinus - feathers	Gift from J. Hudon Royal Alberta Museum	N/A
Cardinalis cardinalis - feathers	Gift from C. Brown at University of Tulsa	N/A
Serinus canaria forma *domestica*	Lopes et al.[21]	N/A
Chemicals, peptides, and recombinant proteins		
Papain	Worthington Biochemical Corporation	LS003126
Dulbecco's Modified Eagle Media	Thermo Fisher Scientific	11995-040
Fetal bovine serum	Thermo Fisher Scientific	10437-028
DNase1	Roche	04716728001
Bovine Serum Albumin	Invitrogen	AM2618
RNase inhibitor	New England Biolabs	M0314S
Trizol LS reagent	Thermo Fisher Scientific	10296028
RNeasy Mini kit	Qiagen Sciences	74104
GRS cDNA Synthesis Kit	GRiSP	GK80.0100
Superscript IV reverse transcriptase	Invitrogen	18090050
Power SYBR Green PCR Master Mix	Applied Biosystems	4368577
DIG RNA Labeling Mix	Roche	11277073910
Polyethylenimine (MW 25000)	Polysciences	23966-1
Hexane	Alfa Aser	45652
tert-butyl methyl ether	Arcos	177040025
Acetonitrile	Fisher	A998
Methanol	Fisher	A452
Dichloromethane	Alfa Aser	22917
Isopropanol	Alfa Aser	22906
Tetrahydrofuran	Burdick Jackson	340
Trimethylamine	Millipore Sigma	TX1202-5
EasySpin Genomic DNA Kit	Citomed	SP-DT- 250
TruSeq DNA PCR-free Library Preparation Kit	Illumina	20015962
PCR-cDNA barcoding kit	Oxford Nanopore	SQK-PCB109
Tween 40	TCI America	T2531
β-carotene	DSM Nutritional Products	0489999004
Canthaxanthin	DSM Nutritional Products	5005256004
Lutein (Floraglo)	DSM Nutritional Products	5011868022
Zeaxanthin (Optisharp)	DSM Nutritional Products	5003563004
3S,3′S-zeaxanthin	CaroteNature GmbH	Nr. 0121
Astaxanthin (CAROPHYLL Pink)	DSM Nutritional Products	5006589329

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MS222	Sigma-Aldrich	E10521
Epinephrine hydrochloride	Sigma-Aldrich	E4642
Critical commercial assays		
Bicinchoninic acid protein assay kit	Thermo Fisher Scientific	23227
HDL and LDL Calibrator kit	ISE S.r.l.	R0300000009
HDL and LDL Control (reference) kit	ISE S.r.l.	R0400000300
TAG Multicalibrator kit	ISE S.r.l.	R0300000006
HDL assay kit	ISE S.r.l.	R0200001301
LDL Miura 200 assay kit	ISE S.r.l.	R0200001401
TAG Miura 200 assay kit	ISE S.r.l.	R0100000901
Deposited data		
Sorted cone photoreceptor RNA-seq data	NCBI Short Read Archive	PRJNA817619
RNA-seq data from regenerating feather follicles, whole-genome sequencing data and Nanopore long-read sequencing from red-throated parrotfinches	NCBI Short Read Archive	PRJNA764058
Whole-genome sequencing data from diamond firetails and star finches	NCBI Short Read Archive	PRJNA780378
Raw HPLC chromatograms and peak spectra, modified chicken transcriptome file, and tissue carotenoid and lipid data	Dryad	https://doi.org/10.5061/dryad.xsj3tx9j9
Experimental models: Cell lines		
HEK 293 cells	ATCC	CRL-1573
Experimental models: Organisms/strains		
Danio rerio	David Parichy Laboratory, University of Virginia	Wild type ABb
Gallus gallus / white leghorn	Charles River Laboratory	SPF research eggs
Erythrura psittacea	Licensed breeders in Portugal	N/A
Stagonopleura guttata	Licensed breeders in Portugal	N/A
Bathilda ruficauda	Licensed breeders in Portugal	N/A
Oligonucleotides		
Oligonucleotides are listed in Table S6	N/A	N/A
Recombinant DNA		
pME_bdh1a-2A-nEOS	This paper	N/A
pME_cyp2ae2-2A-nEOS	This paper	N/A
pCAG-dsRed-2A-GFP	Toomey et al.[41]	N/A
pCAG-GFP-2A-DsRed	Toomey et al.[41]	N/A
pCAG-CYP2J19-2A-GFP	This paper	N/A
pCAG-RDH5-2A-GFP	This paper	N/A
pCAG-BDH1L-2A-DsRed	This paper	N/A
pCAG-TTC39B-2A-GFP	This paper	N/A
pCAG-cyp2ae2-2A-GFP	This paper	N/A
pCAG-bdh1a-2A-DsRed	This paper	N/A
hsp70l::bdh1a-2A-nEOS	This paper	N/A
hsp70l::cyp2ae2-2A-nEOS	This paper	N/A
Software and algorithms		
Kallisto (v.0.46.2)	Bray et al.[52] and Melsted et al.[53,54]	https://pachterlab.github.io/kallisto/
Sleuth (v.0.30.0)	Pimentel et al.[27]	https://pachterlab.github.io/sleuth/

(*Continued on next page*)

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Trimmomatic (v0.39)	Bolger et al.[55]	http://www.usadellab.org/cms/index.php?page=trimmomatic
HISAT2 (v2.2.1)	Kim et al.[56]	http://daehwankimlab.github.io/hisat2/
Rsubread (v1.22.2)	Liao et al.[57]	https://bioconductor.org/packages/release/bioc/html/Rsubread.html
DESeq2	Love et al.[58]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
limma	Ritchie et al.[59]	https://bioconductor.org/packages/release/bioc/html/limma.html
TransDecoder (v.5.5.0)	N/A	https://github.com/TransDecoder/Transdecoder/releases
MAFFT (v7.487)	Katoh and Standley[60]	https://mafft.cbrc.jp/alignment/software/
AliView (v1.26)	Larsson[61]	https://ormbunkar.se/aliview/
RAxML-NG (v1.0.3)	Kozlov et al.[62]	https://github.com/amkozlov/raxml-ng
ModelTest-NG (v0.1.5)	Darriba et al.[63]	https://github.com/ddarriba/modeltest
trimAl (v1.4.rev22)	Capella-Gutiérrez et al.[64]	http://trimal.cgenomics.org/downloads
FigTree (v1.4.4)	N/A	http://tree.bio.edu.ac.uk/software/figtree/
FastQC (v0.11.8)	N/A	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Picard	N/A	http://broadinstitute.github.io/picard
SAMtools (v1.11)	Li et al.[65]	N/A
Freebayes (v1.3.2)	N/A	https://github.com/ekg/freebayes
PCAngsd (v0.98)	Meisner and Albrechtsen[66]	https://github.com/Rosemeis/pcangsd
prcomp (v3.6.3)	N/A	https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/prcomp
ANGSD (v0.929-13-gb5c4df3)	Korneliussen et al.[67]	https://github.com/ANGSD/angsd
vcftools (v0.1.16)	Danecek et al.[68]	http://vcftools.sourceforge.net/
vcflib (v1.0.2)	Garrison et al.[69]	https://github.com/vcflib/vcflib
Beagle (v5.1)	Browning et al.[70]	https://faculty.washington.edu/browning/beagle/b5_1.html
GEMMA (v0.98.1)	Zhou and Stephens[71]	https://github.com/genetics-statistics/GEMMA
Ensembl (*Biomart* - release 104)	Kinsella et al.[72]	N/A
LUMPY (v0.3.1)	Layer et al.[73]	https://github.com/arq5x/lumpy-sv
NGMLR (v0.2.7)	Sedlazeck et al.[74]	https://github.com/philres/ngmlr
Sniffles (v1.0.12B)	Sedlazeck et al.[74]	https://github.com/fritzsedlazeck/Sniffles
YASS web server	Noé and Kucherov[75]	https://bioinfo.lifl.fr/yass/yass.php
minimap2 (*v2.22*)	Li[76]	https://github.com/lh3/minimap2
NCBI BLAST+ (*v2.12.0*)	Camacho et al.[77]	https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/
Other		
YMC carotenoid column	YMC	CT99S05-2546WT
Chiralpak IA-3 column	Chiral Technologies	80525
Pirkle L-leucine chiral column	Regis Technologies	1-731054-300

RESOURCE AVAILABILITY

Lead contact

Please direct requests for further information, resources, and reagents to the lead contact, Joseph Corbo (jcorbo@wustl.edu).

Materials availability

All unique/stable reagents generated in this study are available from the originating labs with a completed Materials Transfer Agreement. The lead contact will facilitate fulfillment of requests.

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

Data and code availability

- RNA and DNA sequence data have been deposited in the NCBI Short Read Archive and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- Other forms of data including raw HPLC chromatograms and peak spectra, the modified chicken transcriptome file, and tissue carotenoid and lipid measurement values are available in a Dryad repository (DOI listed in the key resources table).
- This paper does not report original code.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell line

Enzymatic assays were performed in HEK 293 cells (which were originally derived from a female embryo) and cultured following the distributor's recommendations (ATCC, CRL-1573).

Animals

Fertilized specific-pathogen-free white leghorn chicken (*Gallus gallus*) eggs (Charles River Laboratories) were incubated at 38°C until hatching. The male and female chicks were then harvested immediately (P0) or reared for 15 days and then harvested (P15). Chicks were reared in groups of 3–5 individuals under constant illumination from a heat lamp and were provided Purina Start & Grow medicated feed (Purina Animal Nutrition LLC) and tap water *ad libitum*. For tissue collection, chicks were euthanized by carbon dioxide inhalation followed by manual cervical dislocation. Chicken husbandry and experimental procedures were carried out in accordance with United States National Institutes of Health guidelines[78] and approved by the Washington University in St. Louis Animal Care and Use Committee (protocol #19-1110).

Adult male and female Red-throated parrotfinches (*Erythrura psittacea*), diamond firetails (*Stagonopleura guttata*), and star finches (*Bathilda ruficauda*) were obtained from aviaries of licensed breeders in Portugal and were kept in indoor cages (1.0 × 0.5 × 0.5 m) with *ad libitum* access to water and food (Prestige tropical finches mix, Versele-Laga, Deinze, Belgium). The orange morph of the red-throated parrotfinch is commonly referred to as 'sea green' in reference to the hue of its green feathers. For tissue collection, birds were euthanized by cervical dislocation. All procedures were carried out in accordance with national and international regulations for the maintenance of live birds in captivity (FELASA, Federation of European Laboratory Animal Science Associations). The procedures were also conducted in accordance with the Directive 2010/63/EU on the protection of animals and were approved by the Animal Welfare and Ethics Body of CIBIO/InBIO (February 2018).

Male and female zebrafish (*Danio rerio*) and pearl danio (*Danio albolineatus*) were reared under standard conditions (~28°C; 14-hour-light/10-hour-dark cycle). Larvae were initially fed marine rotifers derived from high-density cultures and enriched with Rotimac and Algamac (Reed Mariculture). Older larvae and adults were fed live brine shrimp and a blend of flake foods enriched with dried spirulina. Fish were euthanized by immersion in a solution of tricaine methanesulfonate (MS222; 250 mg/l) for 15 min. Fish studies were performed in accordance with United States National Institutes of Health guidelines.[78] The fish studies were approved by the University of Virginia's Animal Care and Use Committee (protocol #4170).

Feathers for carotenoid analysis were obtained from several sources. Rump feathers of European serin (*Serinus serinus*) and breast feathers of European greenfinch (*Chloris chloris*) were donated by licensed breeders from birds raised in captivity. Breast feathers from an adult male northern cardinal (*Cardinalis cardinalis*) were obtained from a specimen in the University of Tulsa collection (USGS permit #20948 to C. Brown). Pine siskin (*Spinus pinus*) feathers were provided by Jocelyn Hudon from museum specimens collected under the auspices of Canadian Wildlife Service permit 20-AB-SC002 (Royal Alberta Museum). Feathers of yellow and 'red factor' canaries (*Serinus canaria* forma *domestica*; referred to simply as '*Serinus canaria*' in the manuscript) were collected as part of an earlier study.[21]

METHOD DETAILS

Fluorescence-activated cell sorting (FACS)

Chicken retinas were removed from the eye by dissection, transferred to calcium and magnesium-free Hank's balanced salt solution (Thermo Fisher Scientific) and dissociated into single cells by papain digestion.[79] A single chick retina was incubated in 800 μl of HBSS with 1.3 mg of papain (Worthington Biochemical Corporation) at 37°C for 10 min. The retina was then dissociated by gentle trituration with a pipette. The dissociated cells were washed with Dulbecco's Modified Eagle Media (DMEM; Thermo Fisher Scientific) containing 10% fetal bovine serum (Thermo Fisher Scientific) and DNase1 (Roche) for 5 min at 37°C. The cells were then pelleted by centrifugation at 1000 RCF for one minute, the media was removed, and the cells were resuspended in 800 μl of sorting buffer (2.5 mM EDTA, 25 mM HEPES, 1% Bovine Serum Albumin in HBSS) containing 100 units of RNase inhibitor (New England Biolabs, M0314S). The cell suspension was filtered twice through a 40 μm cell strainer, and the cells were then subjected to FACS.

Dissociated retinal cells were sorted on a FACS Aria-II (BD Biosciences) operated by the Flow Cytometry Core in the Department of Pathology and Immunology at Washington University in St. Louis. Cells were first gated by forward and side scatter to exclude debris and incompletely dissociated cells. To distinguish the unique oil droplet fluorescence of the individual cone photoreceptor subtypes the following excitation lasers and emission filters were used: FITC – blue laser (488 nm excitation) with a 505 nm long-pass and

530/20 nm band-pass emission filter; PE-Texas Red – yellow/green laser (561 nm excitation) with a 600 nm long-pass and 610/20 nm band-pass emission filter; AmCyan-A – violet laser (405 nm excitation) with a 505 nm long-pass and 610/20 nm band-pass emission filter. Cell populations were then gated based on clustering on FITC vs. PE-Texas Red or FITC vs. AmCyan-A signals and sorted cells were collected into sorting buffer with added RNase inhibitor (described above). At the conclusion of a run, Trizol LS reagent (Thermo Fisher Scientific) was added to the sorted cells and RNA was extracted according to manufacturer's instructions. Alternatively, the samples were stored in Trizol for up to one week at -80°C prior to RNA extraction.

RNA-seq analyses

To perform RNA-seq analysis of chicken cone photoreceptors, total RNA derived from sorted cone subtypes was submitted to the Genome Technology Access Center (GTAC) at Washington University in St. Louis for library preparation and sequencing. The P0 sample was subjected to 150-bp paired-end sequencing on an Illumina NovaSeq 6000; the P15 sample were subjected to 50-bp single-end sequencing on an Illumina HiSeq3000. For P0 samples, a total of 477,375,380 paired-end reads with an average of 31,825,025 reads per individual (range = 23,105,612-49,635,587) were generated (Table S4). For P15 samples, a total of 558,522,622 single-end reads with an average of 37,234,841 reads per individual (range = 26,803,768- 52,970,538) were generated (Table S4). Raw sequences reads in FASTQ format were pseudo-aligned to the updated chicken transcriptome (described below), and read counts were quantified using *Kallisto* (v.0.46.2) with 100 bootstrap samples (parameter –b 100).[52–54] For quantification of P15 data (single-end reads), an estimated fragment length of 200 bp with standard deviation of 20 was used. For P0 data (paired-end reads), the fragment length and standard deviation were calculated automatically by *Kallisto*.

To perform RNA-seq analysis of regenerating feather follicles of red-throated parrotfinches, feather follicle regeneration was induced by plucking small patches of throat feathers from four wild-type and four orange males five days prior to skin excision. To harvest tissue, birds were euthanized by manual cervical dislocation, and skin patches were excised, snap-frozen in liquid nitrogen, and stored at -80°C until RNA extraction. Due to logistical constraints associated with the purchasing of the birds, these experimental procedures were performed on two different occasions, three months apart (October 2019 and January 2020). In each batch, we sampled two wild-type and two orange birds. Total RNA was isolated from the eight samples using the RNeasy Mini kit (Qiagen) followed by DNAse digestion. RNA integrity was measured using a TapeStation RNA ScreenTape (Agilent), and RNA concentration was measured using a Qubit RNA BR assay kit. cDNA was generated from approximately 2.5 μg of RNA using the GRS cDNA Synthesis Kit (GRiSP), according to the manufacturer's instructions. Strand-specific Illumina libraries were prepared using the LM-Seq protocol[80] and sequenced using 150 bp paired-end reads. A total of 640,699,020 paired-end reads with an average of 80,087,378 read per individual (range = 53,260,824-110,246,188) were generated (Table S4).

To compare gene expression between wild-type and orange parrotfinches, raw sequence reads were trimmed of adapters, low quality bases (Phred-score < 5), and shorter sequences (<30 bp) using *Trimmomatic* (v0.39).[55] Since there is no publicly available refence genome for the red-throated parrotfinch, the remaining sequencing reads were mapped to the Gouldian finch (*Chloebia gouldiae*) reference genome using *HISAT2* (v2.2.1).[56] Information on exon and splice-site location was extracted from the annotation file and included in the indexing of the genome to improve computation and mapping sensitivity. For each individual and gene, raw counts were summarized with *featureCounts*, a function of the *Rsubread* package (v1.22.2).[57]

Chicken transcriptome annotation

The most recent chicken transcriptome annotation (*Gallus gallus* annotation release 105) was supplemented with the following genes: *OPN1SW* (NM_205438.1), *RCVRN* (NM_205514.1), *GRK1* (NM_204695.1), *EGR2* (XM_015288137.1), *LOC121112497* (XM_040694195.1), *HOXC4* (XM_015300328.1), *LOC121109415* (XM_040661464.1, XM_040661462.1, XM_040661467.1, XM_040661463.1, XM_040661465.1, XM_040661466.1), *GATA1* (NM_205464.1), *LOC100858452* (XM_015273943.2), *HOOK2* (XM_025145771.1), *HOXC6* (XM_015300352.2), *ACTL6B* (XM_015274812.2), *ZNF281* (XM_015290371.1), *SUV39H1* (XM_025145798.1), *ATXN7L3* (XM_015273210.2, XM_015273211.2, XM_015273212.2), *LOC121112239* (XM_040693044.1), *OCT2* (EF410002.1), *HOXC5* (XM_015300327.2), *ZNF653* (XM_015302690.2), *LOC121112231* (XM_040693016.1), *LOC121112233* (XM_040693018.1), *LOC101747537* (XM_015285687.2), *FOXQ2* (Table S5), and chicken mitochondrial genes derived from NCBI *Gallus gallus* mitochondrial genome sequence and annotation (NC_053523.1). Several gene symbols were corrected in accordance with published literature: *GNG11* to *GNGT1*, *FCN3* to *FIBCD1*, *LOC101748254* to *BDH1L*, and *TKTL1* to *TKT*.

Quantitative real-time PCR analysis

Reverse transcription-quantitative real-time PCR (qPCR) was performed using total RNA extracted from cone subtypes isolated by FACS. Superscript IV reverse transcriptase (Invitrogen) was used to synthesize first-strand cDNAs from RNA, according to the manufacturer's instructions. The qPCR analysis was then performed in three biological replicates using Power SYBR Green PCR Master Mix (Applied Biosystems) on a QuantStudio 3 Real-Time PCR System (Applied Biosystems). All qPCR primers were intron-spanning with the exception of those for *TAL2* (primer sequences in Table S6). Transcript levels of individual genes were normalized to glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) levels.

In situ hybridization

To perform *in situ* hybridization analysis, a canary *BDH1L* template was generated by PCR (primer sequences in Table S6), and sense and anti-sense digoxigenin-labeled probes were produced from this template using established methods.[26,81] Histologic sections of

skin with regenerating feather follicles from yellow and 'red factor' canaries were generated in a prior study.[21] *In situ* hybridization was carried out on these sections as previously described.[26,81] Stained sections were imaged with an Olympus BX-51 microscope.

Phylogenetic analysis of protein sequences

To construct a phylogenetic tree of BDH1L-related proteins, a *TBLASTN* search (version 2.11.0+)[77] was performed using chicken BDH1L (XP_004935295.3) to query the RefSeq transcripts of the following species: zebra finch (*Taeniopygia guttata*), green sea turtle (*Chelonia mydas*), West Indian Ocean coelacanth (*Latimeria chalumnae*), green anole (*Anolis carolinensis*), chicken (*Gallus gallus*), red-eared slider (*Trachemys scripta elegans*), zebrafish (*Danio rerio*), medaka (*Oryzias latipes*), spotted gar (*Lepisosteus oculatus*), mouse (*Mus musculus*), and human (*Homo sapiens*). All sequences with E-values $< 1 \times 10^{-10}$ were retained for the initial analysis. Pearl danio (*Danio albolineatus*) genes orthologous to zebrafish *bdh1*, *bdh1a* (*zgc: 113142*), and *rdh5* were also included in the analysis. Retrieved cDNA sequences were translated into amino acid sequences using *TransDecoder.LongOrfs* in *TransDecoder* (v.5.5.0). The resultant sequences were aligned using a multiple sequence alignment program (L-INS-i) in *MAFFT* (v7.487)[60] under default settings. Identical sequences with different names, sequences with species-specific significant gaps in data, sequences with poor alignment in regions of strong homology, and partial sequences were removed, and members of BDH1, BDH1L, and RDH5 sub-families were then manually selected in an alignment viewer and editor *AliView* (v1.26).[61] The aligned sequences were trimmed with *trimAl* (v1.4.rev22)[64] using the '-automated1' option, and gapped sites were removed using the '-nogaps' option. The maximum likelihood tree was inferred by *RAxML-NG* (v1.0.3).[62] The best tree was selected out of twenty alternative runs on ten random and ten parsimony-based starting trees (–tree pars[10], rand[10] option). The amino acid replacement models of Le-Gascuel (LG) with gamma distribution (G4) and proportion of invariant sites (I) were selected using the Akaike information criterion implemented in *ModelTest-NG* (version 0.1.5).[63] The bootstrap values were obtained by sampling 1,000 times. The results were viewed using *FigTree* (version 1.4.4).

To construct a phylogenetic tree of CYP2J19-related proteins, a similar approach was used. *TBLASTN* searches were performed using chicken CYP2J19 (XP_422553.3) and zebrafish Cyp2ae2 (XP_017209221.1) to query the RefSeq transcripts of the following species: sea lamprey (*Petromyzon marinus*), Australian ghost shark (*Callorhinchus milii*), chicken, zebrafish, and human. All sequences with E-values $< 1 \times 10^{-10}$ were retained for the initial analysis. Also included in the analysis were pearl danio genes orthologous to zebrafish *cyp2ae1* and *cyp2ae2* and genes in zebra finch, Chinese softshell turtle (*Pelodiscus sinensis*), and painted turtle (*Chrysemys picta bellii*) orthologous to chicken *CYP2J19*. After translation, alignment, manual selection, and re-alignment, the final aligned sequences were trimmed with *trimAl* using the '-automated1' option. The maximum likelihood tree for one hundred protein sequences was inferred using the same methods and parameters described in the previous paragraph. Le-Gascuel (LG) with gamma distribution (G4) and proportion of invariant sites (I) was used as the amino acid replacement model.

Cell-based assay of enzymatic activity

To perform enzymatic assays, full-length versions of the most abundant transcript isoforms of *RDH5, CYP2J19*, and *TTC39B* were obtained by PCR using cDNA derived from total RNA extracted from whole retina of P0 white-leghorn chickens (protein sequences in Table S5; primer sequences in Table S6). A full-length transcript of chicken *BDH1L* could not be obtained by PCR, possibly due to the presence of high GC content. Instead, a codon-optimized version of the coding sequence was synthesized (GeneScript) based on XM_004935238.3 (Table S5). The coding sequences of *D. albolineatus bdh1a* and *cyp2ae2* were subcloned from the trangenesis constructs (pME_*bdh1a-2A-nEOS* and pME_*cyp2ae2-2A-nEOS*) described in a subsequent section. For the enzymatic assay, coding sequences were subcloned into the first position of a bicistronic expression construct that encoded a fluorescent protein in the second position (*pCAG-[first position]-2A-GFP or pCAG-[first position]-2A-dsRed*). The sequence of all expression constructs was confirmed by Sanger sequencing.

Enzymatic assays were performed in cultured HEK 293 cells (ATCC, CRL-1573) grown to 80% confluency and then transiently transfected with one or more expression constructs using polyethylenimine (Polysciences). Control cultures were transfected with a bicistronic construct containing fluorescent proteins in both positions (*pCAG-dsRed-2A-GFP*). Transfected cells were incubated for 16-24 hours, and construct expression was confirmed by visualizing fluorescent protein expression. The all-*trans* isomers of carotenoid substrates (zeaxanthin, lutein, and β-carotene) were purified by HPLC fractionation from samples provided by DSM (zeaxanthin was purified from Optisharp; lutein from Floraglo; β-carotene from 'β-carotene 10%'). ε,ε-carotenoid substrates (canary feather peaks 1 and 2) were extracted from canary feathers as described below and purified by HPLC. The product of CYP2J19 metabolism of zeaxanthin (Figure 3A Peak 2) was collected from HEK 293 cells expressing CYP2J19 and TTC39B, extracted as described below, and purified by HPLC. For each assay, the substrates were added individually to the cell culture media at concentrations ranging from 0.14 to 1.0 μg/mL in a solution of 0.035% Tween 40 (TCI).

To determine the stereoselectivity of the CYP2J19/BDH1L-mediated conversion of zeaxanthin into astaxanthin, enzymatic assays were also performed using the following substrates: 3R,3′R-zeaxanthin (purified from Optisharp), 3R,3′S-zeaxanthin (purified from marigold petals), 3S,3′S-zeaxanthin (purchased from Carotenature), or an equimolar mixture of all three stereoisomers. The cultured cells were incubated for an additional 16-24 hours with the carotenoid substrate. After incubation, the media was removed, and the cells were washed twice with phosphate-buffered saline (PBS) before resuspension in 1 mL PBS for subsequent HPLC analysis.

Carotenoid extraction

To extract carotenoids for HPLC analysis, the following procedures were used.

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

Cultured cells

Cultured cells and media were scraped from the culture dish and pelleted by centrifugation. The media was then removed, and the cells were resuspended in 500 μL of 0.9% NaCl containing 0.1 g of zirconia beads. The cells were then disrupted using a Beadbug homogenizer (Benchmark Sci) for 30 sec at 4 kHz. 250 μl of 100% ethanol was then added to the solution, and the solution was mixed by vortexing. 500 μl of hexane:*tert*-butyl methyl ether (1:1, vol:vol), was subsequently added to the solution which was ground once more for 30 sec at 4 kHz. The resultant cell homogenate was centrifuged at 10,000 RCF for 3 mins, and the upper solvent fraction (containing carotenoids) was collected in a 2 mL glass vial. The hexane:*tert*-butyl methyl ether extraction was repeated two more times, all extracts from an individual sample were combined, and the solvent was then completely evaporated under a stream of nitrogen.

Feathers

The pigmented barbs were trimmed from the feathers, weighed on an analytical balance, then washed in 100% ethanol, followed by 100% hexane, and then air-dried overnight. The washed feather barbs were then ground in 1mL of methanol containing 0.1 g of zirconia beads on a Beadbug homogenizer (Benchmark Sci.) for 15 mins at 4 kHz. The homogenate was then centrifuged, and the solvent collected and dried under a stream of nitrogen. Extraction was repeated 2–5 times until the remaining solid feather material appeared colorless.

Fins

Three caudal fin samples from each transgenic line or control zebrafish were pooled and extracted as described above for cultured cells. To de-esterify fin carotenoids, the dried extracts were saponified by resuspension in 0.02 M NaOH in methanol, capped under nitrogen, and incubated for 4 hours at room temperature in the dark. Carotenoids were then re-extracted with hexane:*tert*-butyl methyl ether (1:1, vol:vol) and dried under a stream of nitrogen.

Retina

Carotenoids were extracted individually from whole retinas (two retinas/bird) of three wild-type and three *orange* mutant red-throated parrotfinches by grinding in solvent as described above for cultured cells. To de-esterify retinal carotenoids, each extract was split into two subsamples and saponified with weak base (0.02M NaOH in methanol) to recover ketocarotenoids or strong base (0.02M NaOH in methanol) to recover other carotenoid types following the procedures described above for fins.

Plasma

Plasma (20 μL) from three wild-type and three *orange* mutant red-throated parrotfinches were extracted individually by adding 250 μL of 100% ethanol followed by 250 μL of hexane:*tert*-butyl methyl ether (1:1, vol:vol) and vortexing to mix. The samples were then centrifuged, and the supernatant was collected and dried under a stream of nitrogen.

Liver

Liver tissue (~5 μL fragment) from three wild-type and three *orange* mutant red-throated parrotfinches were extracted individually as described above for cultured cells. Prior to the addition of solvent to the samples, an aliquot of the liver homogenate was set aside for measurement of total protein content using a bicinchoninic acid (BCA) assay (Thermo Fisher Scientific).

High-performance liquid chromatography

To perform reverse-phase HPLC analysis, carotenoids were dissolved in acetonitrile:methanol (1:1, vol:vol) and injected into an Agilent 1200 series HPLC fitted with a YMC carotenoid column (5.0 μm, 4.6 mm × 250 mm). Samples were then eluted with a gradient mobile phase of acetonitrile:methanol:dichloromethane (44:44:12, vol:vol:vol) through 11 min, ramping up to acetonitrile:methanol:dichloromethane (35:35:30) from 11-21 min, followed by isocratic conditions through 35 min. Solvent was pumped at a constant rate of 1.2 mL/min, and the column was maintained at 30°C. Sample elution was monitored using a UV-Vis photodiode array detector at various wavelengths, and carotenoids were identified by comparison to authentic standards of astaxanthin, canthaxanthin, zeaxanthin, lutein, and β-carotene (DSM) or by comparison to carotenoids extracted from feathers whose carotenoid content is known.

For stereochemical analysis, zeaxanthin stereoisomers were fractioned and collected on a Daicel Chiralpak IA-3 column (Chiral Technologies)[82] with a mobile phase of hexane:isopropanol (90:10, vol:vol), at a flow rate of 0.5 mL/min, and the column temperature set to 25°C. Astaxanthin products were fractionated and collected on the reverse-phase HPLC system (as described above) and then analyzed on a Pirkle L-leucine chiral column (Regis Technologies)[83] with a mobile phase of hexane:tetrahydrofuran:trimethylamine:acetonitrile (77:19:2:2, vol:vol:vol:vol), at a flow rate of 1.5 mL/min, and the column temperature set to 25°C. Stereoisomers of astaxanthin were identified by comparison to a racemic mixture of synthetic astaxanthin (DSM) and 3S,3'S-astaxanthin (extracted from *Haematococcus pluvialis*[83]).

Mass spectrometry

To perform mass spectrometry analysis, carotenoid products of cell culture-based enzymatic assays (described above) were fractionated by HPLC and dried under nitrogen. Extracts were then dissolved in *tert*-butyl methyl ether:methanol (1:1), and 20 μL was injected into an Agilent 1290 ultra-high performance liquid chromatography instrument (UHPLC) fitted with a YMC carotenoid column (5.0 μm, 4.6 mm × 250 mm). The flow rate was 0.5 mL/min, and the column temperature was held at 40°C. Samples were separated using variable proportions of solvent A (methanol:water 80:20 [vol:vol] with 0.1% formic acid) and solvent B (*tert*-butyl methyl ether:methanol:water, 78:20:2 [vol:vol] with 0.1% formic acid). The gradient was as follows: beginning at 80%:20% (solvent A:solvent B), increasing linearly to 90% solvent B over 15 min, then increasing linearly to 99% solvent B over 1 min, holding at 99% B for 6 min, and then immediately returning and holding at 20% B for 3 min. The UHPLC eluent was detected using in-line diode-array-detection

at 450 and 468 nm, followed by ionization and analysis using a quadrupole time-of-flight mass spectrometer (Agilent 6545 LC/Q-TOF). The atmosphere pressure chemical ionization (APCI) source was operated in positive ion mode, using the following parameters: gas temperature = 350°C, gas flow = 5 L/min, nebulizer = 60 psig, vaporizer = 400°C, voltage capillary = +3500 V, and corona = 4 μA. MS scan experiments collected data over 100-1700 *m/z*, with an acquisition rate of 3.0 (spectra/sec). Iterative MS/MS and targeted MS/MS fragmentation were also performed, using fixed collision energies at 10 and 25 eV, scanning for products between 100-1700 *m/z*, at 3 spectra/sec, using narrow isolation width (1.3 *m/z*). The carotenoid products were compared to authentic standards of astaxanthin, zeaxanthin, lutein and fractionated yellow canary feather extracts (peaks 1 and 2 in Figure 4).

Transgene expression in zebrafish

To ectopically express pearl danio (*Danio albolineatus*) genes in zebrafish, we first amplified the full-length coding sequences of pearl danio *bdh1a* and *cyp2ae2* (protein sequences in Table S5) using cDNA reverse-transcribed derived from total RNA extracted from whole anal fin tissue of adult fish (primer sequences in Table S6). These amplicons were cloned into the pJET vector (Thermo Fisher Scientific) and then subcloned by Gibson assembly to create pME_*bdh1a-2A-nEOS* or pME_*cyp2ae2-2A-nEOS*. These bicistronic cassettes were then placed under the control of the *hsp70l* promoter using Gateway cloning (Tol2Kit)[84] to create the final constructs for zebrafish transgenesis: *hsp70l*::*bdh1a*-2A-*nEOS* and *hsp70l*::*cyp2ae2*-2A-*nEOS*.

To generate transgenic zebrafish, *hsp70l::bdh1a-2A-nEOS* and *hsp70l::cyp2ae2-2A-nEOS* plasmids were injected, individually and in combination, into ~600 zebrafish embryos at the one-cell stage along with *tol2* mRNA. Fish were raised to the juvenile stage (10 mm length) and then heat shocked for 1 hour at 38°C, 6 times/day for 10 consecutive days. Six injected fish with strong, widespread fluorescent protein expression were selected for analysis along with six control fish (randomly selected from an uninjected group). The sexes are indistinguishable at this stage and fish were sampled randomly with respect to sex. Fish were imaged using a Zeiss AxioZoom stereomicroscope (v16) equipped with an Axiocam 506 camera. Caudal fins were then harvested, placed in 1x PBS, and stored at -80°C for later carotenoid analysis.

Retina imaging

To visualize retinal oil droplet mosaics, birds were sacrificed by manual cervical dislocation. The eye was then removed from the body, the anterior segment excised, and the eye cup fixed in 4% paraformaldehyde (PFA) for 5 min at room temperature, followed by 10 min at 37°C. The fixed retinas were then removed from the eye cup and dissected free from the retinal pigment epithelium in calcium and magnesium-free HBSS buffer. The retina was then rinsed three times in the same buffer. The retina was then mounted photoreceptor side up on an uncoated microscope slide in a small pool of buffer, and a coverslip was placed on top. The oil droplet mosaic was visualized using a Leica DM6000B microscope. Images were captured with a Leica MC170 HD camera and processed using V4.13 Leica Application Suite software.

WHOLE-GENOME-SEQUENCING AND READ MAPPING

To identify the mutations causing yellow/orange feather coloration in red-throated parrotfinches (*Erythrura psittacea*), diamond firetails (*Stagonopleura guttata*), and star finches (*Bathilda ruficauda*), genome-wide polymorphism data were generated by whole-genome sequencing. A total of 107 samples were collected from captive birds that belonged to multiple breeders in Portugal (Table S3). For a given species, individual birds were sampled from multiple breeders to reduce the likelihood of close relatedness among individuals. Whole blood was drawn from live birds into a heparin-free capillary tube by brachial venipuncture with a sterile needle and stored in 96% ethanol for later DNA extraction. DNA was isolated from these samples using an EasySpin Genomic DNA Kit SP-DT-250 (Citomed), including an RNAse A digestion step. DNA quality and purity were assessed using spectrophotometry (Nanodrop) and fluorometric quantitation (Qubit dsDNA BR Assay Kit, Thermo Scientific). High-quality DNA was then used to generate sequencing libraries for 107 individual birds. Red-throated parrotfinch samples were prepared for sequencing using the TruSeq DNA PCR-free Library Preparation Kit (Illumina). The Diamond firetail and Starfinch samples were prepared using a modified version of Illumina's Nextera XT protocol.[85] Libraries were sequenced using 150-bp paired-end reads on an Illumina instrument at Novogene to an average coverage of 13.4X, 1.1X, and 1.1X, for the red-throated parrotfinch, diamond firetail, and starfinch, respectively (Table S3).

After sequencing, read quality was inspected with *FastQC* (v0.11.8). The sequencing reads were mapped to the reference genome of the closest species with a published annotated genome: the red-throated parrotfinch to the Gouldian finch (*Chloebia gouldiae*) genome,[86] and the diamond firetail and starfinch to the zebra finch genome (taeGut3.2.4).[87] The reads were mapped with *bwa-mem* (v0.7.17-r1188)[88] using default parameters, and duplicates were identified and removed using *Picard MarkDuplicates*. Overall sequencing and mapping summary statistics were computed using *SAMtools* (v1.11)[65] (Table S3).

Whole-genome Nanopore sequencing

Structural rearrangements in the orange morph of the red-throated parrotfinch were also evaluated using long-read sequencing. Whole-genome sequencing data was generated for one orange male using Oxford Nanopore long-read Technology. High Molecular Weight (HMW) DNA was extracted from whole-blood stored in ethanol using a modified salt-based protocol.[89] To prevent ethanol contamination, blood was washed in PBS overnight prior to HMW DNA isolation. Purified genomic DNA (gDNA) concentration and integrity were assessed using a NanoDrop spectrophotometer, Qubit dsDNA BR Assay Kit, and TapeStation gDNA

ScreenTape (Agilent). Two independent unsheared libraries were prepared from 8 μg of DNA using the SQK-LSK109 kit and sequenced on a MinION. Libraries were run separately on a FLO-MIN106 flow cell with a nuclease flush step in between runs (according to NL_9076_v109 flow cell cleanup protocol). Fast5 files were base-called with *Guppy* (v4.3.4; Oxford Nanopore software), and the resulting reads were concatenated for subsequent analyses. A total of 929,016 reads (average length = 9.3 Kb) were analyzed, resulting in approximately ~7.9X coverage of the genome.

Genetic crossing of parrotfinches

To investigate the inheritance of the *orange* mutation in the red-throated parrotfinch and confirm the mutation's association with the Z-chromosome (see results), controlled crosses between orange and red individuals were performed. In total, 20 pairs were crossed. To evaluate sex-linked inheritance, 14 of the 20 crosses involved males and females of different phenotypes in all possible combinations. From these crosses, a total of 88 individual offspring were obtained; the phenotype (orange or red) was scored after the offspring fledged.

Sequencing of full-length transcripts using Nanopore technology

To evaluate differential *TTC39B* transcript isoform expression in retina and developing feather follicles, Oxford Nanopore sequencing technology was used. Feather follicle regeneration was induced by plucking small patches of throat feathers eight days prior to follicle harvest. Regenerating follicles and retinas were sampled from two homozygous wild-type and two *orange* mutant male parrotfinches. Harvested follicles and retinas were stored in *RNAlater* buffer at -20°. Total RNA was extracted from regenerating follicles and retinas as described above, and 50 ng of total RNA from each sample was used for library preparation. A total of eight DNA libraries were prepared using the PCR-cDNA barcoding kit (SQK-PCB109) protocol, with the following PCR program: initial denaturation at 95°C for 30 secs followed by 14 cycles of denaturation at 95°C for 15 secs, annealing at 62°C for 15 seconds and extension at 65°C for 20 min (allowing for transcripts up to ~20 Kb), and final extension at 65°C for 6 min. The multiplexed libraries were sequenced on a MinION instrument (FLO-MIN106 flow cell). Base-calling was carried out in *Guppy* using the fast configuration (dna_r9.4.1_450bps_fast.cfg). Barcode trimming was enabled to detect mid-strand barcodes (as expected from chimeric reads) with the minimum score set at 65. Reads were concatenated based on barcode ID for subsequent analyses. A total of 4,113,270 reads (averaging ~514,159 reads per library) were analyzed (Table S4).

The resulting reads were mapped to the zebra finch reference genome (bTaeGut1_v1.p) using *minimap2* (*v2.22*)[76] with appropriate parameters for Oxford Nanopore Technology cDNA reads. The zebra finch genome is the most comprehensively annotated reference genome of any passerine bird;[87] it thus provides a more complete set of described isoforms for the target gene. Aligned reads were converted to bed12 format and corrected for misaligned splice sites using the genomic annotation. *TTC39B* isoform structure was visualized in a genome browser and compared between wild-type and *orange* mutant parrotfinches for both tissues. To further evaluate potential effects of the *TTC39B* duplication on isoform expression, exons from known isoforms of zebra finch *TTC39B* were obtained from *Ensembl* (release 104) and a BLAST database was constructed using *NCBI BLAST+* (*v2.12.0*).[77] All reads that mapped to the *TTC39B* genomic region were then compared to the resulting database using BLAST, and the exons were characterized across the transcripts. This allowed the identification of duplicated exons and the detection of other aberrant transcript isoforms.

Lipoprotein and fatty acid analyses

To perform lipoprotein and fatty acid analyses of red-throated parrotfinches, blood, liver, and fat were sampled from four red and four orange adult male individuals. Prior to tissue harvest, all birds were kept between six to eight weeks in our facilities under standard conditions (described above). To obtain liver extracts, liver samples stored at -80°C were placed in buffer (50 mM Tris-HCl pH 7.5, 4 mM EDTA acid, 50 mM NaF, 0.5 mM phenylmethylsulfonyl fluoride, 1 mM 1,4-dithiothreitol and 250 mM sucrose [1:5, w:v]) and homogenized using an IKA T10 Standard Turrax homogeneizer (IKA-Werke GmbH & Co. KG, Germany).[90] Homogenates were centrifuged (15,000 RCF; 40 min; 4°C), and the supernatant was then transferred to Miura 200 vial (I.S.E. S.r.l.; Guidonia, Italy). Blood was collected from the brachial vein into a heparinized capillary tube pierced with a 26-gauge needle (Microvette CB 300; Sarstedt AG & Co, Germany). Plasma was subsequently separated by centrifugation (10000 RCF; 10 min) and stored at -20°C. The sample was later thawed on ice and transferred to Miura 200 vials for analysis.

Plasma and liver HDL, LDL, and triacylglycerol (TAG) levels were quantified using a fully automated Miura 200 analyzer with associated reagent kits (I.S.E. S.r.l.; Guidonia, Italy; A-R0100000901, A-R0200001301, and A-R0200001401). The analyses were performed in two biological replicates. Triacylglyceride (A-R0300000006) and HDL and LDL (A-R0400000300) were run against a reference sample for quality control.

For fatty acid (FA) analysis, lipids were extracted from pectoral fat (previously stored at -80°C) using a solution of methyl *tert*-butyl ether, methanol, and water (MTBE:MeOH:H_2O, 10:3:2.5; 20 mL/g of tissue).[91] The extract was then transferred to amber-colored glass vials, dried, and stored at -20°C. TAG samples were reconstituted in 250 μL of deuterated chloroform ($CDCl_3$, ACROS ORGANICS, code: 166252500, Lot: A0397242, 99.8 atom % D) containing a pyrazine standard. Proton Nuclear Magnetic Resonance (^{1}H NMR) spectra were obtained using a Varian VNMRS 600 MHz NMR (Agilent, Santa Clara, CA, USA) spectrometer equipped with a 3 mm broadband probe with z-gradient. Fully relaxed ^{1}H NMR spectra of TAG samples were acquired with a sweep width of 8 KHz, a 90° pulse and 8 s of recycling time (3 s of acquisition time and 5 s pulse delay). A total of 16 FIDs were acquired for each ^{1}H NMR spectrum. Spectra were analyzed with ACD/NMR Processor Academic Edition software (ACD/Labs, Advanced Chemistry Development). The FA profile (in percentage) was estimated by ^{1}H NMR because lipid species such as the omega-3 fatty acids (n-3 FA; 0.90

ppm), monounsaturated fatty acids (MUFA; 1.90 ppm) and polyunsaturated fatty acids (PUFA; 2.00 and 2.30 ppm) provide distinguishable peaks in specific regions of the spectrum, while the SFA were calculated by difference.[92] A methodological control for TAG extraction was performed, and the FA/glycerol ratio was calculated from the area of all FA α protons times two, divided by TAG-bound glycerol sn1,3 protons. The FA/glycerol ratio was consistent and ~3 for all birds, as expected for pure TAG preparations.

QUANTIFICATION AND STATISTICAL ANALYSIS

Unless otherwise specified statistical analyses were carried out in R.[93] Statistical significance was defined as a p value < 0.05 with adjustments for multiple comparisons as detailed below.

RNA-seq statistical analyses

Sleuth (v.0.30.0)[27] was used to compare transcript counts among the five cone photoreceptor subtypes; this analysis was performed separately for P0 and P15 data. Very lowly expressed transcripts were excluded by the *Sleuth* default filter that requires a minimum of 5 reads in at least 47% of biological replicates for a transcript to be included in the analysis. Differential expression among cone subtypes was determined with a likelihood ratio test comparing a model with the single factor of 'cone subtype' to a reduced model that assumed equal expression across all cone subtypes. Transcripts were pseudo-aligned, quantified, and analyzed individually by the algorithm. To determine differential expression at the gene level, the p values for all transcripts of each gene were aggregated as previously described.[94] A list of the top differentially expressed genes ('top hits' in Data S1) was generated by selecting the genes with a q-value $\leq$ 0.05 and a mean expression value $\geq$ 5 transcripts per million (TPM) reads in at least one cone subtype. A custom script was used to categorize each gene according to its pattern of differential enrichment among cone subtypes. A gene was defined as 'enriched' in one or more cone subtypes if its expression value in that subtype (or subtypes) was $\geq$ 2-fold higher than in the next highest-expressing subtype. Secondary enrichment categories were defined in a similar fashion. Differentially expressed genes which failed to show $\geq$ 2-fold expression differences between any two subtypes were categorized as 'complex'.

Differential gene expression between the regenerating feather follicles of wild-type and orange parrotfinches was determined using *DESeq2*,[58] with a significance threshold of FDR = 0.1. To control for potential batch effects resulting from individuals being sampled on different dates, 'batch' was included as a variable in the model testing for differential expression. When required for graphical representations, normalized read counts were also corrected for potential batch effects using the function *removeBatchEffect* implemented in *limma*.[59]

Quantitative real-time PCR analysis

Differential gene expression among cone subtypes was determined by one-way ANOVA and post-hoc contrasts compared with Tukey's honestly significant difference (HSD) test. Additional statistical details can be found in Table S1.

Quantitation of red-throated parrotfinch tissue carotenoid, lipoprotein, and fatty acid content

The carotenoid content of the feathers, retinas, plasma, and livers of red and orange morph red-throated parrotfinches were compared with a Welch's two-sample t test. The lipoprotein and fatty acid composition between the liver and plasma of red and orange morph red-throated parrotfinches were compared with Mann-Whitney U tests. Additional statistical details are provided in the legend accompanying Figure S7.

Quantitation of ketocarotenoid production in cell culture

To determine if co-expression of TTC39B significantly enhanced ketocarotenoid production by CYP2J19 and BDH1L, the mass of ketocarotenoids as a proportion of total carotenoids in the cells was calculated for each replicate culture and compared between conditions with a Welch's two-sample t test. Additional statistical details are provided in the legend accompanying Figure 7I.

Variant calling and functional annotation

To identify the mutations causing yellow/orange feather coloration in red-throated parrotfinches (*Erythrura psittacea*), diamond firetails (*Stagonopleura guttata*), and star finches (*Bathilda ruficauda*), genome-wide polymorphism data were generated by whole-genome sequencing (details above). The Bayesian haplotype-based method implemented in *Freebayes* (v1.3.2) was used for variant calling. Reads with a mapping quality < 30 or an individual base quality < 20 were excluded from the analysis. Positions with less than three alternative alleles and an overall coverage two times higher than the average were also excluded. The latter filter is expected to remove regions of the genome with copy number variation and/or regions that are highly repetitive. The '*genotype-qualities*' option was also used to obtain a quality score associated with each individual genotype. The variants (both SNPs and indels) retained after filtering were functionally annotated using the genetic variant annotation and effect prediction toolbox *SnpEff* (v4.3t).[95]

Genome-wide association and genetic differentiation analyses

Prior to genome-wide association analysis, patterns of population structure among the sequenced individuals of the three species were investigated using Principal Component Analysis (PCA) as implemented in *PCAngsd* (v0.98).[66] This software was used to generate a covariance matrix between all individuals based on estimates of allele frequency derived from genotype likelihoods that was subsequently used to estimate principal components and individual loadings with the R function *prcomp* (v3.6.3). This

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

analysis was not based on the variant and genotype calls of *Freebayes*, but on an independent analysis based on genotype likelihoods generated using *ANGSD* (v0.929-13-gb5c4df3).[67]

A genome-wide association study was performed for red-throated parrotfinch based on individual genotypes. The genotype of each individual was coded as 'missing data' if its quality was below 20 ("–minGQ 20" in *vcftools*), or its coverage was below 5X or higher than 46X (i.e., twice the average coverage of the individual with the highest coverage; "–minDP 5 –maxDP 46" in *vcftools*). Following genotype filtering but prior to the analysis, several stringent filtering steps were applied to the catalog of variants. Using *vcftools* (v0.1.16),[68] the following variants were removed from the analysis: i) indels and all non-biallelic markers; ii) variants with a quality score below 500 ("–minQ 500" in *vcftools*); iii) variants having more than one third of individual genotypes missing ("–max-missing-count 30" in *vcftools*); iv) variants with a minor allele frequency less than 10% ("–maf 0.1" in *vcftools*); and v) variants that deviated from Hardy-Weinberg equilibrium ("–hwe 0.001" in *vcftools*). Using *vcffilter* (*vcflib* v1.0.2),[69] additional variants were removed from the analysis based on the following criteria: i) allele balance ("AB > 0.25 & AB < 0.75 | AB < 0.01"); ii) discrepancy in the properly paired status of reads supporting reference or alternate alleles ("PAIRED > 0.05 & PAIREDR > 0.05 & PAIREDR / PAIRED < 1.75 & PAIREDR / PAIRED > 0.25 | PAIRED < 0.05 & PAIREDR < 0.05"); and iii) alleles that are only seen on one strand ("SAF > 0 & SAR > 0 & RPR > 1 & RPL > 1").

These filtering steps yielded a total of 3,514,701 SNPs for the genome-wide association analysis. Using this catalog of variants, missing genotypes were imputed with the software *Beagle* (v5.1)[70] using default parameters. Association analyses were then performed with *GEMMA* (v0.98.1)[71] coding plumage color as a case-control (red versus orange) and using a univariate linear mixed model. To account for potential confounding effects of relatedness among individual birds, a kinship matrix was constructed using the "-gk" option in *GEMMA*, which was used as a random effect in the linear mixed model. p values were calculated using a likelihood ratio test (LRT). A Bonferroni correction threshold was used to determine SNPs significantly associated with the color phenotype ($-\log_{10}$ [0.05/3,514,701] = 7.85).

Due to lower sequence coverage and the smaller number of individuals in our genomic analyses of diamond firetail and starfinch, the region associated with red and orange coloration was mapped by estimating genetic differentiation across the genome using a sliding window approach. Genetic differentiation was summarized using the fixation index (F_{ST}), which was estimated by means of genotype probability methods as implemented in *ANGSD* (v0.929-13-gb5c4df3).[67] Reads with a mapping quality < 30 or with multiple mapping coordinates, as well as individual bases within each read with quality < 20, were excluded from the analysis. Windows with less than 50% of the positions passing filters were also excluded. F_{ST} was averaged in windows of 50 Kb moved in steps of 10 Kb along each scaffold.

Identity-by-descent mapping (IBD)

For the red-throated parrotfinch, identity-by-descent mapping was used to refine the position of the candidate gene within the associated region. This approach relies on the assumption that a given mutation arose only once and then increased in frequency. Accordingly, the causative gene or mutation is expected to be contained within the minimum common haplotype shared among individuals homozygous for the mutant allele. This approach is expected to work well with the *orange* mutation, which appeared during domestication and is recessive. To identify candidate regions for the orange phenotype as a complement to the association analysis (see section above), the genome was first screened for signatures of reduced genetic diversity. To control for variation in genetic diversity across the genome that exists even in wild-type individuals, the ratio of nucleotide diversity (π) in wild-type red parrotfinches and orange parrotfinches ($\pi_{\text{wild-type}}/\pi_{\text{orange}}$) was obtained using a sliding window approach (50 Kb windows with 10 Kb steps). This ratio is expected to be elevated around the genomic region underlying the orange phenotype. Nucleotide diversity was calculated using genotype probability methods as implemented in *ANGSD* (v0.929-13-gb5c4df3).[67] Reads with a mapping quality < 30 or with multiple mapping coordinates, as well as individual bases within each read with quality < 20, were excluded from the analysis. Windows with less than 75% of the positions passing filters were also excluded.

To examine patterns of haplotype similarity and homozygosity in the candidate region identified by the association and genetic diversity analyses (see main manuscript), the individual genotypes described above were used. This analysis was restricted to male birds, which carry two copies of the Z-chromosome and can therefore be heterozygous for a given position. A total of 15 red birds and 10 orange birds were considered. For each position, each individual was coded as 'homozygous' or 'heterozygous' and this information was plotted across the associated region. The *Ensembl* (release 104) *Biomart* online tool[72] was used to extract the gene content of the candidate interval that exhibited high homozygosity in orange individuals.

Detection of structural rearrangements

The candidate region for the red-throated parrotfinch was initially screened for structural rearrangements using Illumina short-read whole-genome resequencing data from 45 individuals (see above). Two different approaches were used. First, *LUMPY* (v0.3.1)[73] was used with default parameters. This program uses a combination of multiple signals in the data, including paired-end alignment, split-read alignment, and read-depth information to detect structural rearrangements. Second, read depth was evaluated by calculating the number of reads in nonoverlapping 1 Kb windows across the genome of each individual. To detect potential changes in coverage due to duplications or deletions, the number of reads for a given window was divided by the genome-wide average. In the absence of copy number variation, most windows are expected to have values close to 1. The individual values were then averaged among all orange and red individuals, and read depth was plotted and compared between the two phenotypes. Candidate structural variants inferred from both approaches were manually inspected and intersected with the Gouldian finch genome annotation.

Please cite this article in press as: Toomey et al., A mechanism for red coloration in vertebrates, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.08.013

To evaluate for the occurrence of structural variants underlying the orange parrotfinch morph, nanopore long-reads were mapped to the Gouldian finch reference genome using *NGMLR* (v0.2.7)[74] with parameter presets for Oxford Nanopore Technology data. We then used *Sniffles* (v1.0.12b),[74] a structural variant caller specific to third-generation sequencing data, to detect large (>10 bp) structural variants against the Gouldian finch scaffold (QUSF01000011.1: 8,000,000–9,000,000 bp).

In parallel, raw Oxford Nanopore Technology reads were assembled into a *de novo* genome assembly using *wtdbg2* software[96] with default parameters and assuming an estimated genome size of 1.1 Gb. A final consensus sequence was derived with the *wtpoa-cns* function. The draft assembly had a total size of ~952 Mb (3,562 contigs) and a N50 of ~653 Kb. To better visualize sequence similarity and structural variation, the full contig containing *TTC39B* (805,841 bp) was retrieved and compared to the homologous Gouldian finch scaffold by means of a dotplot in the *YASS* web server.[75]