Red Carotenoid Coloration in the Zebra Finch Is Controlled by a Cytochrome P450 Gene Cluster

Graphical Abstract

Highlights
- The yellowbeak mutation maps to a narrow region of chromosome 8 with a CYP cluster
- CYP2J19 loci most likely encode ketolases that generate red ketocarotenoids
- CYP2J19 loci are involved both in red coloration and red retinal oil droplets
- Involvement of cytochrome P450s provides a novel mechanism of signal honesty

Authors
Nicholas I. Mundy, Jessica Stapley, Clair Bennison, ..., Tim R. Birkhead, Staffan Andersson, Jon Slate

Correspondence
nim21@cam.ac.uk (N.I.M.), staffan.andersson@bioenv.gu.se (S.A.)

In Brief
Mundy et al. have identified genes required for the bright-red coloration that many birds use for communication, such as attracting mates. They uncover a genetic connection between red coloration and color vision in birds and propose that redness may be an honest signal of mate quality by indicating a bird’s ability to detoxify harmful substances.

Accession Numbers
KX024636
KX024637
KX024638
KX184728
KX184729
Red Carotenoid Coloration in the Zebra Finch Is Controlled by a Cytochrome P450 Gene Cluster

Nicholas I. Mundy,1,4,* Jessica Stapley,2,4 Clair Bennison,2 Rachel Tucker,2 Hanlu Twyman,1 Kang-Wook Kim,2 Terry Burke,2 Tim R. Birkhead,2 Staffan Andersson,3,* and Jon Slate2

1Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK
2Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK
3Department of Biological and Environmental Sciences, University of Gothenburg, Göteborg 40530, Sweden
4Co-first author

*Correspondence: nim21@cam.ac.uk (N.I.M.), staffan.andersson@bioenv.gu.se (S.A.)

http://dx.doi.org/10.1016/j.cub.2016.04.047

SUMMARY

Bright-red colors in vertebrates are commonly involved in sexual, social, and interspecific signaling [1–8] and are largely produced by ketocarotenoid pigments. In land birds, ketocarotenoids such as astaxanthin are usually metabolically derived via ketolation of dietary yellow carotenoids [9, 10]. However, the molecular basis of this gene-environment mechanism has remained obscure. Here we use the yellowbeak mutation in the zebra finch (Taeniopygia guttata) to investigate the genetic basis of red coloration. Wild-type ketocarotenoid genes were absent in the beak and tarsus of yellowbeak birds. The yellowbeak mutation mapped to chromosome 8, close to a cluster of cytochrome P450 loci (CYP2J2-like) that are candidates for carotenoid ketolases. The wild-type zebra finch genome was found to have three intact genes in this cluster: CYP2J19A, CYP2J19B, and CYP2J40. In yellowbeak, there are multiple mutations: loss of a complete CYP2J19 gene, a modified remaining CYP2J19 gene (CYP2J19a), and a non-synonymous SNP in CYP2J40. In wild-type birds, CYP2J19 loci are expressed in ketocarotenoid-containing tissues: CYP2J19A only in the retina and CYP2J19B in the beak and tarsus and to a variable extent in the retina. In contrast, expression of CYP2J19B is barely detectable in the beak of yellowbeak birds. CYP2J40 has broad tissue expression and shows no differences between wild-type and yellowbeak. Our results indicate that CYP2J19 genes are strong candidates for the carotenoid ketolase and imply that ketolation occurs in the integument in zebra finches. Since cytochrome P450 enzymes include key detoxification enzymes, our results raise the intriguing possibility that red coloration may be an honest signal of detoxification ability.

RESULTS AND DISCUSSION

Bright-red coloration of skin or plumage is a major component of avian color diversity and depends, with few exceptions, on the ability to enzymatically convert yellow dietary carotenoids (xanthophylls) to red C4 ketocarotenoid pigments such as astaxanthin or canthaxanthin. Carotenoid ketolase is thus an important key innovation in avian evolution and color diversification [11–13], and many well-known examples of sexually selected visual displays involve bright-red ketocarotenoid coloration in the plumage or beak, functioning to attract partners or to deter rivals (e.g., [5–8]). The anatomical location of ketolation has been contentious. The absence of ketocarotenoids in the blood plasma of some birds suggests peripheral action, close to or in the ketocarotenoid-pigmented tissues, such as the beak or feather follicles [14]. In other species, high concentrations of ketocarotenoids in the plasma imply transport to the periphery from the tissue where ketolation occurs [15, 16], which is likely to be the liver [17].

Ketocarotenoids also occur in avian retinas, where astaxanthin is the pigment in the red oil droplets [18] that are present in longwave-sensitive (LWS) cones and act as cut-off filters to improve color discrimination [19]. The mechanism behind retinal astaxanthin, and its relationship with integumentary ketocarotenoid pigmentation is unknown.

Male zebra finches have ketocarotenoid-based red bill coloration, which is a classical sexually selected trait, with females preferring males with redder bills [6]. Males also have red tarsi, and females have less saturated red bills and tarsi, containing the same ketocarotenoids as the males at lower concentrations [18]. The yellowbeak mutation, only described from captive birds, has yellow bills and tarsi in both sexes and is considered an autosomal recessive mutation among cage-bird enthusiasts (Figure 1) [20].

Ketocarotenoids Are Absent in the Beak of yellowbeak Birds

High-performance liquid chromatography (HPLC) analysis of carotenoids extracted from the beak of yellowbeak birds revealed the presence of dietary yellow carotenoids and an absence of the four red ketocarotenoids (astaxanthin, canthaxanthin, z-doradexanthen, and adonirubin) present in wild-type birds (Figure 2). This is consistent with a role for the yellowbeak mutation in the metabolic conversion of dietary carotenoids to ketocarotenoids.

Yellowbeak Maps to a Narrow Region of Chromosome 8 Containing Candidate Ketolase Genes

We first mapped the yellowbeak mutation using 352 genome-wide SNPs in a pedigree of 95 birds. A total of 17 (0.28) of the second-generation birds expressed the yellowbeak phenotype, almost...
exactly the 3:1 ratio of wild-type-yellowbeak phenotypes expected for a segregating autosomal Mendelian recessive trait. The yellowbeak trait was genetically linked to a single SNP on chromosome 8 (Tgu8) (two-point linkage of yellowbeak to Tgu_SNPS_01464; recombination fraction = 0 cM; logarithm of the odds (LOD) = 5.72; position of SNP on assembly = 25.44 Mb) (Figure 3A).

We next performed fine-scale mapping in this region, obtaining data from 39 SNPs between 15.02 and 27.73 Mb on chromosome 8. Six consecutive markers, positioned between 24.01 Mb and 25.35 Mb, produced recombination fractions of zero with yellowbeak (Table S2; Figure 3B). The flanking markers with some recombination to yellowbeak are located at 22.85 Mb and 25.68 Mb, and so yellowbeak must lie within this 2.83 Mb interval. In fact, the fine mapping provides strong support for a much narrower interval. At the central four of the r = 0 markers, the F2 yellowbeak birds and F0 yellowbeak founders had identical genotypes (Figure S1). The first and sixth markers were the least informative and so had lower power to detect recombination events with yellowbeak (Table S2). The two markers with highest LOD scores were rs83092174 (LOD score = 7.53), which is located at 24,442,889 bp on chromosome 8, and rs83092728 (LOD score = 6.92), which is located at 24,628,232 bp. The two flanking markers had weaker support (LOD = 5.42 in both cases), and the interval spanning the flanking markers covers 24,269–25,029 Mb. This interval contains five genes on the zebra finch genome assembly, two CYP2J-like loci, NF1A, HOOK, and FGGY (Figure 3C). Thus, there is greatest statistical support for yellowbeak to reside within this 0.76 Mb window, but because some F1 parents are not informative at all SNPs, an extended window of errors in the current zebra finch genome assembly (Tae-gut3.2.4) in this region. The existing assembly indicates two adjacent CYP2J-like loci on chromosome 8, with a further two CYP2J-like loci on an unassembled chromosome 8 contig. Using a combination of long-range PCR and Illumina MiSeq sequencing, we found direct evidence for three adjacent CYP2J-like loci on chromosome 8, incorporating some sequence from the previously unanchored region (Figure 3D; Supplemental Experimental Procedures; Figure S1B). Two of these loci, which we term CYP2J19A and CYP2J19B, are close homologs related to chicken CYP2J19, whereas the third, CYP2J40, is more distantly related.

Long-range PCR results indicated the presence of an ~13 kb deletion of a full copy of CYP2J19 in yellowbeak birds (Figure S1D). Using a PCR assay (Figure S1C), we found a perfect association between this deletion and yellowbeak (Fisher’s exact test of association between phenotype and absence of the wild-type allele, p = 1.16 × 10⁻⁰⁵): all yellowbeak birds were homozygous for the deletion (N = 12), whereas wild-type birds were homozygous (N = 9) or heterozygous (N = 9) for the intact allele. The single copy of CYP2J19 in yellowbeak birds, CYP2J19⁰, is divergent from both wild-type CYP2J19 loci and has four unique amino acid substitutions. Phylogenetic reconstruction shows that it is closer to CYP2J19B than CYP2J19A, implying that the mutation may have involved a full deletion of CYP2J19A (Figure S2A). Substantial allelic variation in wild-type alleles of CYP2J19B, but not CYP2J19A, was detected, with two divergent allelic lineages present (Figure S2A). A 76 bp indel polymorphism in CYP2J19B only was found ~600 bp upstream of the start codon. An intact copy of the other CYP2J-like locus on chromosome 8, CYP2J40, is present in yellowbeak, and this encodes a protein with a single amino acid substitution (Y209N) compared to wild-type CYP2J40.

CYP2J19 Loci Are Specifically Expressed in Ketocarotenoid-Containing Tissues, and Bill Expression Is Greatly Reduced in yellowbeak

In wild-type birds, we detected expression of CYP2J19 loci in the beak, tarsus, and retina, which are all sites of ketocarotenoid deposition, but not in the heart, or liver and duodenum, which
are important sites for carotenoid metabolism and uptake (Figure S2B). CYP2J19A expression was detected only in the retina, whereas CYP2J19B was strongly expressed in the beak and tarsus, with variable expression in the retina (Figure S2C). In contrast, CYP2J40 was expressed in all of these tissues (data not shown).

In qRT-PCR experiments, CYP2J19yb expression was barely detectable in the bill of yellowbeak birds (Figure 4A) and was significantly lower than expression of CYP2J19B in wild-type birds (two-tailed t test, p < 0.05). In contrast, CYP2J40 expression was present in yellowbeak retinas and similar to that of CYP2J19A/B in wild-type retinas (Figure 4B). There were no differences in CYP2J40 expression between yellowbeak and wild-type bills (Figures 4C and 4D). Taking genetic and expression data together, our results suggest that the deletion in yellowbeak removed both CYP2J19A and cis-regulatory sequences driving expression of CYP2J19B in the beak, bringing CYP2J19yb under control of cis-regulatory sequences of CYP2J19A, thereby driving expression of CYP2J19yb in the retina.

The bills of female zebra finches have the same ketocarotenoids as males, but at lower concentrations. In an experiment to investigate potential sex-biased expression, we did not find significant differences in CYP2J19B expression in the bills of wild-type males and females (Figure S2D).

Taken together, our results strongly suggest that the CYP2J2-like cluster is required for red ketocarotenoid coloration. The pattern of expression of CYP2J19 loci strongly implicates a role of CYP2J19B in red beak and tarsus pigmentation and suggests that CYP2J19A is probably involved in astaxanthin production in the retina, potentially in combination with CYP2J19B. Since CYP2J40 has a fixed mutational difference in yellowbeak, we cannot rule out an additional role of this locus in coloration.
the majority of CYP2J19 in ketocarotenoid presence in those tissues. The absence of
vides strong circumstantial evidence that they play a direct role
ketocarotenoids are deposited (beak, tarsus, and retina) pro-
1438
Current Biology
act as detoxification enzymes. Our results therefore provide
coloration [24, 25].

Overall, the CYP2J2-like loci are strong candidates to encode
the long-sought ketolase enzyme. These results concur with an
independent study that has used introgression of red coloration
into the canary to identify CYP2J19 as a putative ketolase [23].
Direct experiments on the biochemical properties of the
enzymes will be required to confirm this. Our results further
demonstrate the utility of mutations in domesticated forms for
determining specific expression in the tissues where
ketocarotenoids are deposited. Further, they suggest an
intriguing genetic link between ketocarotenoids used for red
coloration and those used for color vision. The presence of red
oil droplets is thought to be conserved across birds [26], whereas
red coloration is patchily distributed [13], so a role of CYP2J2-like
loci in color vision is likely to have evolved prior to a function in
coloration.

Ancestry of yellowbeak and CYP2J19 Duplication
The presence of a derived indel polymorphism that only occurs on
the yellowbeak haplotype and multiple fixed mutations in the
CYP2J2-like cluster in yellowbeak together imply that yellow-
beak is an old mutation that probably predates domestication. In
a preliminary analysis to determine whether the yellowbeak allele
is still segregating in the wild, we genotyped 50 wild zebra
finches from a wild population and did not detect the yellowbeak
deletion (data not shown). The apparent origin of the yellowbeak
allele in the wild together with the presence of divergent alleles of
CYP2J19B segregating in wild-type zebra finches suggests
interesting evolutionary dynamics at this locus, perhaps
including selection, that warrant further study. The contribution
of other genetic variation at CYP2J19B to beak color variation
is another interesting question. A quantitative trait locus study
of beak redness in captive zebra finches did not find evidence
for a quantitative trait locus on chromosome 8, although sugges-
tive linkage to loci associated with beak color was detected on
four other chromosomes [27].

Blast searches of available avian genomes, published studies
on avian CYPs [28, 29] and our unpublished data from weaver-
birds (Ploceidae) (H.T., unpublished data), which are closely
related to estrildids, indicate that most other birds, from
chickens to canaries, have a single CYP2J2 locus. Hence, the
duplication of CYP2J19 in zebra finches into copies with their
own tissue-specific expression is relatively recent. It will be inter-
esting to determine when the duplication occurred in the avian
phylogeny and whether it was associated with a change in color
signaling.

Implications for Sexual Selection
A key issue in sexual selection research is the maintenance of
honest signaling [30]. For carotenoid-based displays, much of
the extensive literature has focused on the resource allocation
hypothesis, the notion that it may be costly to irreversibly divert
carotenoids to coloration from other physiological roles such as
immune defense and antioxidant functions [31–33]. In other hy-
potheses, such as the “shared-pathway hypothesis,” some or
all carotenoids may not themselves be limiting resources, but
their uptake and metabolism, in particular ketocarotenoid deri-
vation, may be intrinsically linked to crucial aspects of phenotypic
and/or genetic quality, such as vitamin A homeostasis [34, 35]
(but see 36]) or respiratory efficiency [37]. Our finding that
cytochrome P450s are involved in carotenoid metabolism brings
an exciting new perspective to this field. The cytochrome P450
oxidase enzymes, present in all living organisms, are primary
metabolizers of foreign substances. In vertebrates, they are
best known as detoxification enzymes [38], sharing several
basic requirements for their function, including protein cofactors.

Figure 4. Quantitative Real-Time RT-PCR Results in Wild-Type
and yellowbeak Birds
Expression is normalized relative to β-actin. Error bars indicate the SEM. The
top panel shows results for CYP2J19A, CYP2J19B, and CYP2J19C in the beak (A)
and retina (B). The lower panel shows results for CYP2J40 in in the beak (C)
and retina (D). n = 3 wild-type and n = 3 yellowbeak birds throughout. See also Figure S2.
(cytochrome reductase and/or cytochrome b_{5}) and an optimal cellular redox state. Many cytochrome P450s are expressed in the inner mitochondrial membrane [38], leading to a potential direct association between ketolation and respiration [35]. An alternative and more specific “honest signaling” hypothesis is that CYP2J2-based ketocarotenoid coloration is advertising some aspect of cytochrome P450 functionality in general (tissue-specific or systemic), such as detoxification ability.

**ACCESSION NUMBERS**

The new nucleotide sequences reported here have been submitted to GenBank under accession numbers KX024636-KX024638, KX184726, and KX184729.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Supplemental Experimental Procedures, two figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2016.04.047.

**AUTHOR CONTRIBUTIONS**

N.I.M., J. Stapley, J. Slate, S.A., and T.R.B. designed the experiments; N.I.M., J. Stapley, S.A., C.B., R.T., and H.T. obtained the data; N.I.M., J. Stapley, J. Slate, S.A., and K.-W.K. analyzed the data; N.I.M., S.A., J. Stapley, and J. Slate drafted the manuscript; and all authors verified the final version of the manuscript.

**ACKNOWLEDGMENTS**

We thank the BBSRC (grant numbers BB/E017509/1 and BB/I02185X/1 to J. Slate, T.B., and T.R.B.), the EC Marie Curie International Outgoing Fellowship scheme (project 253300 to J. Stapley), the NERC Biomolecular Analysis Facility (NBAF961 to N.I.M. and J. Slate), Murray Edwards College, Cambridge (to N.I.M.), and the Swedish Research Council (to S.A.) for funding. We thank Simon Griffith for wild zebra finch samples.

Received: September 17, 2015
Revised: March 18, 2016
Accepted: April 14, 2016
Published: May 19, 2016

**REFERENCES**

1. Foote, C.J., Brown, G.S., and Hawnyshyn, C.W. (2004). Female colour and male choice in sockeye salmon: implications for the phenotypic convergence of anodromous and nonanadromous morphs. Anim. Behav. 67, 69–83.
2. Hill, G.E., and McGraw, K.J., eds. (2006). Bird Coloration, Volume 2, Function and Evolution (Harvard University Press).
3. Khan, S.A., Levine, W.J., Dobson, S.D., and Kralik, J.D. (2011). Red signals in male zebra finch: evidence for the inner mitochondrial membrane [38], leading to a potential direct association between ketolation and respiration [35]. An alternative and more specific “honest signaling” hypothesis is that CYP2J2-based ketocarotenoid coloration is advertising some aspect of cytochrome P450 functionality in general (tissue-specific or systemic), such as detoxification ability.

1. Foote, C.J., Brown, G.S., and Hawnyshyn, C.W. (2004). Female colour and male choice in sockeye salmon: implications for the phenotypic convergence of anodromous and nonanadromous morphs. Anim. Behav. 67, 69–83.
2. Hill, G.E., and McGraw, K.J., eds. (2006). Bird Coloration, Volume 2, Function and Evolution (Harvard University Press).
3. Khan, S.A., Levine, W.J., Dobson, S.D., and Kralik, J.D. (2011). Red signals in male zebra finch: evidence for the potential direct association between ketolation and respiration [35]. An alternative and more specific “honest signaling” hypothesis is that CYP2J2-based ketocarotenoid coloration is advertising some aspect of cytochrome P450 functionality in general (tissue-specific or systemic), such as detoxification ability.

1. Foote, C.J., Brown, G.S., and Hawnyshyn, C.W. (2004). Female colour and male choice in sockeye salmon: implications for the phenotypic convergence of anodromous and nonanadromous morphs. Anim. Behav. 67, 69–83.
2. Hill, G.E., and McGraw, K.J., eds. (2006). Bird Coloration, Volume 2, Function and Evolution (Harvard University Press).
3. Khan, S.A., Levine, W.J., Dobson, S.D., and Kralik, J.D. (2011). Red signals in male zebra finch: evidence for the potential direct association between ketolation and respiration [35]. An alternative and more specific “honest signaling” hypothesis is that CYP2J2-based ketocarotenoid coloration is advertising some aspect of cytochrome P450 functionality in general (tissue-specific or systemic), such as detoxification ability.
common cormorant: evolutionary relationships with 130 diapsid CYP2 clan sequences and chemical effects on their expression. Comp. Biochem. Physiol. C Toxicol. Pharmacol. 153, 280–289.

29. Watanabe, K.P., Kawai, Y.K., Ikenaka, Y., Kawata, M., Ikushiro, S., Sakaki, T., and Ishizuka, M. (2013). Avian cytochrome P450 (CYP) 1-3 family genes: isoforms, evolutionary relationships, and mRNA expression in chicken liver. PLoS ONE 8, e75689.

30. Andersson, M. (1994). Sexual Selection (Princeton University Press).

31. von Schantz, T., Bensch, S., Grahn, M., Hasselquist, D., and Wittzell, H. (1999). Good genes, oxidative stress and condition-dependent sexual signals. Proc. Biol. Sci. 266, 1–12.

32. Svensson, P.A., and Wong, B.B.M. (2011). Carotenoid-based signals in behavioural ecology: a review. Behaviour 148, 131–189.

33. Simons, M.J.P., Cohen, A.A., and Verhulst, S. (2012). What does carotenoid-dependent coloration tell? Plasma carotenoid level signals immunocompetence and oxidative stress state in birds-A meta-analysis. PLoS ONE 7, e43088.

34. Hill, G.E. (2011). Condition-dependent traits as signals of the functionality of vital cellular processes. Ecol. Lett. 14, 625–634.

35. Johnson, J.D., and Hill, G.E. (2013). Is carotenoid ornamentation linked to the inner mitochondria membrane potential? A hypothesis for the maintenance of signal honesty. Biochimie 95, 436–444.

36. Simons, M.J.P., Groothuis, T.G.G., and Verhulst, S. (2015). An appraisal of how the vitamin A-redox hypothesis can maintain honesty of carotenoid-dependent signals. Ecol. Evol. 5, 224–228.

37. Hill, G.E. (2014). Cellular respiration: the nexus of stress, condition, and ornamentation. Integr. Comp. Biol. 54, 645–657.

38. Danielson, P.B. (2002). The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. Curr. Drug Metab. 3, 561–597.
Supplemental Information

Red Carotenoid Coloration in the Zebra Finch

Is Controlled by a Cytochrome P450 Gene Cluster

Nicholas I. Mundy, Jessica Stapley, Clair Bennison, Rachel Tucker, Hanlu Twyman, Kang-Wook Kim, Terry Burke, Tim R. Birkhead, Staffan Andersson, and Jon Slate
Figure S1: A. Heatmap of genotypes of F2 birds at SNPs around yellowbeak. Four consecutive SNPs show perfect co-segregation with yellowbeak in the F2 birds with the yellowbeak phenotype. Because the typed SNPs are not fixed for alternative alleles between F0 yellowbeak and F0 wild type founders, some wildtype founders also contain alleles seen in yellowbeak birds and a perfect association between phenotype and genotype is not expected (or observed) in the wildtype F2 birds. The four consecutive SNPs are located between 24.269Mbp and 25.028Mbp. The CYP2J2-like cluster is located at ~24.573Mbp. B. Arrangement of CYP2J2-like loci. 1. Current zebra finch assembly (Taegut3.2.4). The region of chromosome 8 shown is the 2-LOD drop interval 24.269-25.029Mbp as in main Figure 3C, containing two CYP2J2-like loci. Unordered chromosome 8 has two further CYP2J2-like loci. 1. Experimentally determined arrangement of CYP2J2-like loci on chromosome 8, shown in opposite orientation to 1. The total length of this segment is ~40kbp. Solid and open boxes indicate unique sequences of 600–800bp that are unlinked on the current assembly but in fact are 13kbp apart at either end of the coding sequence of CYP2J19A. C. Long-range PCR assay for the deletion of a whole CYP2J19 locus. 1-3 represent PCR amplicons with 3 sets of primer pairs (CJ2up9F/CJ1ex1R2, CJ2up9F/CJ2Aex9R and CJ2up4F/CJ1ex1R2, respectively). D. Gel image of long-range PCR showing the ~13kbp deletion present in yellowbeak. Lanes 1 and 7: DNA ladder; lanes 2 and 3: two wildtype birds; lanes 4 and 5: two yellowbeak birds; lane 6: negative control. Figure S1 is related to Figure 3.
Figure S2. A. Maximum likelihood-based phylogenetic reconstruction of a 12kbp alignment of CYP2J19 loci in zebra finch (Tgu) with the single CYP2J19 of the ploceid southern red bishop (Euplectes orix, Eorix) as an outgroup. The two divergent alleles identified at locus CYP2J19B in wildtype birds are labelled CYP2J19B_1 and CYP2J19B_2. The clade containing the two CYP2J19B alleles, and the clade containing these two alleles plus CYP2J19yb, both had 100% bootstrap support. B. Representative qualitative RT-PCR results showing presence of CYP2J19 expression (both CYP2J19A and CYP2J19B amplified together) in beak, retina and tarsus but not other tissues of wildtype male zebra finch. Results for β-actin are a positive control for cDNA integrity. B = beak, R = retina, T = tarsus, L = liver, H = heart, D = duodenum, c = negative control. DNA ladder is a 100bp ladder from 100bp to 1kbp. C. Representative results showing relative amount of CYP2J19A and CYP2J19B expression in beak (B) and retina (R) of four male zebra finches using a RT-PCR-RFLP assay. RT-PCR product with primers CJ2ex1F2/CYP2J2-2R (~550bp) was digested to completion with BsrI, which cuts CYP2J19A but not CYP2J19B. Results show that only CYP2J19B is expressed in beak, CYP2J19A is only expressed in retina, and CYP2J19B has variable expression in retina - from almost no expression (individual 3), to higher expression than CYP2J19A (individual 2). D. Quantitative real-time RT-PCR results for CYP2J19B in the beak of wildtype males (N = 4) and females (N = 4). Expression is normalised relative to β-actin. Error bars are SEM. Figure S2 is related to Figure 4.
### Table S1 PCR primers (all 5’-3’)

Genomic primers for long-range PCR (see Supplementary Figure 1.)

| Primer     | Sequence                  |
|------------|---------------------------|
| CJ2up9F    | CACACCAGGAAGGCTGGAAGA     |
| CJ2up4F    | AAGTGGGATGACACAGGGAGC     |
| CJ2Aex9R   | GCTGAACCTGGTTGATGGGA      |
| CJ1ex1R2   | GGAAGGGGTTGGGAGGGAAATTA   |

cDNA primers for *CYP2J19A* and *B, CYP2J19*[^b]

qRT-PCR primers for all CYP2J19 loci

| Primer     | Sequence                  |
|------------|---------------------------|
| CYP2J2-2F  | CGGCCAAATATTCCACTCCTCC    |
| CYP2J2-2R  | TGGTAGTCAAGCGGTTCCCA      |

RT-PCR primers for full coding sequence

| Primer     | Sequence                  |
|------------|---------------------------|
| CJ2ex1F2   | TCCTGGGATGAAAAACCTGTGAGT  |
| CJ2ex9ER   | GAGAGAATTCGATTCAGCTCTGAGGCT |
| CJ2Aex9R   | GCTGAACCTGGTTGATGGGA      |
| CJ2do1R    | TTCTGCTGCGCTTCCAGGTGAA    |

[^b]: specific for *CYP2J19A*

CJ1ex1F2   GAGGAAGAGCTCAGAACGGGC
CJ1ex9R2   GAGGAAGAGCTCAGAACGGGC

cDNA primers for *β*-actin

| Primer     | Sequence                  |
|------------|---------------------------|
| ACTF2      | CTCCCTGATGCCCAGGT CAT     |
| ACTR2      | TGGATAACCACAGGACTCCAT     |
### Table S2: Summary of 39 SNPs used for fine mapping of yellowbeak

| SNP          | Flank                                                                 | MAF | Chr 8 Position (bp) | LOD  | R     |
|--------------|------------------------------------------------------------------------|-----|---------------------|------|-------|
| rs83065974   | AAAGCCAGCAAGAAGAAGCCTTTGCTTCCTTCA [C/G] AGATGCATTGCAAAGCTTTTGCTGCAATTGACAG | 0.22 | 15,023,874          | 0.36 | 0.300 |
| rs83066642   | AAGCTTTGAGTTCTTTCTTCATTATTCAGAAAGCT [G/T] TATTTCAGGATCTACAGGTTGCTTTTGGGAAAAT | 0.42 | 15,265,787          | 0.29 | 0.286 |
| rs83068104   | ACTGATTGCTCCATGCTCCCTTGGTCTTCACA [C/G] GAACCTGCTGCAACTGTGATTTCAGAAGGTGTTGA | 0.46 | 15,890,782          | 0.05 | 0.444 |
| rs83068368   | CACCTTTGATATCCTTATGCCTTTGAAATGCAGAAT [G/T] AGTTACTAAGATTTGAGAGGTGGGAAACATGAGA | 0.48 | 15,987,518          | 0.93 | 0.292 |
| rs83068818   | AACATAGTTTGTCTTGCTGACGAGACCTTACAC [A/G] TCCTATGTTAAGAAATGCGGATTTTGTGCTCTACAT | 0.23 | 16,135,096          | 2.12 | 0.083 |
| rs83069593   | GAACAGCTACATCCTCCGTTAAGACGTTGGA [A/G] AAGGGCGCTTTCAGGTTGCTCAAACCTCCTACTTCA | 0.24 | 16,379,407          | 1.35 | 0.111 |
| rs83070604   | TCTTGGCTTTGCTTCTTAGCACCACACAGTCTAGG [C/T] GTTTTCAGATAGCTGCAAGATGAAAACAAAGCA | 0.33 | 16,803,712          | 1.26 | 0.200 |
| rs83071063   | GTGCATCAGCACGGGTAATCTGCTACCCGCTTTA [C/T] TGAAGAGATCTTCTGGTGTATATGTGTAAT | 0.39 | 16,975,150          | 0.04 | 0.400 |
| rs83071968   | TTATACGCTAGGCTGACTGCTGCTTGGTTTTTGGACT | 0.44 | 17,398,777          | 1.35 | 0.111 |
| **ZFRAD_47430** | TGGGAGTGGCCTCGTACCTCCTCCGAGTTTTTGGG [A/G] TCCTATACCTTTTGGGATACCAACACCGACGATCCA | 0.31 | 17,600,146          | 0.43 | 0.308 |
| rs83074941   | TTCTTCTTGTCCTACCCACACCACTTACTGACTA [G/T] CAGAGGAAACACTTCTCCGCTCAAATATTG | 0.45 | 18,540,436          | 0.44 | 0.290 |
| rs83075340   | GTGAGGAGTGAGCACTGCTTCGGCAAGGAGAATG | 0.31 | 18,694,816          | 0.86 | 0.231 |
| rs83075630   | CCAGCCAACTACAGCCCATTGCTGTGCTGACCT | 0.19 | 18,772,650          | 2.65 | 0.071 |
| rs83076198   | ACAGGGGAGCTGTGTAACCTGAGCAAGAAATG [A/G] ACTGCTTGGTATCTTCTGGTATATGAAATGG | 0.36 | 18,920,850          | 1.26 | 0.167 |
| rs83078167   | AAAATAGAGGTAAGAAAGCTTCTTTGGTATCTGAGTGA [A/G] CAGACGCAAAACATCTTCCGCTCAAATATTG | 0.29 | 19,552,783          | 0.84 | 0.200 |
| rs83078802   | AGCAAGAAGCCAGTCTCTTGTTGAGGAGAATTG | 0.30 | 19,722,125          | 2.12 | 0.158 |
| rs83081256   | CCAAGGAATCGGACAGCTTGTCCGATCCT | 0.28 | 20,586,922          | 2.35 | 0.150 |
| rs83081627   | AAATCTCTTTAAGCCAGGCTGAGTAAAACCC [A/G] CAGGCCTGCTTTGCTGCTCAAATATTG | 0.45 | 20,716,455          | 1.05 | 0.182 |
| rs83081937   | TAGGGGTCCAGAGCTGCAAGCAGCACAGGTTTA [C/G] TCTCACAGGAAAGCAGAAGAGTCTTCAAAGAAGAGAT | 0.38 | 20,826,223          | 1.26 | 0.167 |
| rs83082459   | CGTCTCTCAGGATTTGATAACTTACGTTGCTTTCTGCAATCTGAGGCT | 0.18 | 21,069,701          | 0.86 | 0.143 |
| rs83083477   | GTGGAGGAAATACCTTGCCGACCACTTGGCTAGCAGA [A/G] TCTGTGAGATCTTGGGAAAGATATGGAGG | 0.44 | 21,455,773          | 0.32 | 0.353 |
| rs83084659   | AATAGGTGGAGCCAGCATGCTGTCAATCCCGCTA | 0.48 | 21,959,624          | 2.09 | 0.182 |
| rs83085361   | CTGGAAGAGTTAACACAGCAGGCTTACGTTAAGAAGGGTTGAGTGGG | 0.40 | 22,222,774          | 1.90 | 0.167 |
| rs83085927   | TTCCACAGCAGCTCATTCCAGAGCAGGTTGATTTCCTC [C/T] GTGAGACTCCTGCTCATATCACAGAATATCTCAGA | 0.27 | 22,438,476          | 2.38 | 0.077 |
| SNP         | Flank                                                                 | MAF | Chr 8 Position (bp) | LOD  | R     |
|-------------|------------------------------------------------------------------------|-----|---------------------|------|-------|
| rs83086529  | TGCAACAACAGTTACATTTAAACAGTAAATCCTCA [C/T] TGCTTGCTCCAAAACAGGGACAGCAAAATATTC | 0.28 | 22,599,500          | 2.41 | 0.000 |
| rs83086968  | AAAATTTAAGACTGTAATCAATCAAAGACCTTG [G/T] GCTAAATCAAAAGCCATGACCTGTTTTATATTACC | 0.49 | 22,751,681          | 2.35 | 0.150 |
| rs83087266  | TCCCCACTGCCTTTCCATCCACGTCCACCCAGAG [A/C] GGAGAGACAAACACTCTGAGTCCAACGTGACAT | 0.17 | 22,854,724          | 0.07 | 0.333 |
| rs83090816  | ATGGTACCAAGAAAATATTTATCAGTGGGAAAGAAATAT [A/G] CAGCAAGGTGTTTTTTTTAAAAAGATGCTGCTTGAAAT | 0.19 | 24,007,670          | 2.41 | 0     |
| ZFRAD_47955 | GGGGAAGTTTACAGGGGCTGCTACGTAATTCA [A/G] ATGGAAACATTACTGCTGAGTACAGTCTATTG | 0.27 | 24,269,239          | 5.42 | 0     |
| rs83092174  | TAAAATCTTGGAATTTGGGAATCTCTATATCAACA [G/T] TGCTCTTGGTTGATTACCTAGGCTGCTGGTCTTTTTC | 0.41 | 24,442,899          | 7.53 | 0     |
| rs83092728  | TCAATGTCGTTTTTGTCTCTGTCCCCCTTGGAAATA [A/G] CACTCTTAAATATTAATTATCACGCTGCTACA | 0.36 | 24,628,232          | 6.92 | 0     |
| rs83093674  | TGTAGGGGGAATAGCAGCACGTGCTATACGAGGGG [C/G] ACTTCTTCTAATCTTAAATTATATCATCAGCTGCTACA | 0.48 | 25,028,749          | 5.42 | 0     |
| rs83094454  | TGCCAAAGGGAAGGTTGCTCTTTGGAATTTGTCTTCA [A/G] TGTTGTTGATCTCTACGCTGCTGCTGAGCATTCT | 0.24 | 25,350,001          | 3.01 | 0     |
| rs83095237  | TGCACATGGAGCAACACAGGGGACAGGAGCAGTACGAGGTCGAGGGGACTTTTCTCTCT | 0.49 | 25,683,402          | 4.86 | 0.045 |
| rs83095935  | ACATCTCTAGAGCTTGGAGCTTGGAGCTTGTCGCCCTGAGG [A/G] AGATTTAGGCTTTTTAAACTGCTCAGTGAAGTT | 0.37 | 25,876,935          | 1.86 | 0.091 |
| rs83099498  | ACGACTGTGCTGCTATCAAGCCACCAAGAAATCAGGAGGCTGCTGACAGGAGGACAGCCAACAAA | 0.44 | 27,458,783          | 0.86 | 0.143 |
| rs83100186  | TTTTATGACACTGATTATTGCTTCACAAGCTT [C/T] CCACACTGTCACTGGGAAAACTGACGTTGGGCCCATT | 0.13 | 27,688,487          | 0.90 | 0     |
| rs83100244  | CTCTCTCCAGAAAAGTAATCTCTGCTCCTCTACC [A/G] GCTTTGGCTCCTTGTGAATAGAAGAAGCACTACATG | 0.32 | 27,728,768          | 0.96 | 0.263 |

MAF = minor allele frequency, LOD = two-point LOD score with yellowbeak, R = recombination fraction with yellowbeak.
Supplemental Experimental Procedures

Zebra finch sampling

Zebra finches were part of a domesticated population maintained at the University of Sheffield since 1985 under the stewardship of TRB and maintained with ad lib food and water, as described in [S1]. The yellowbeak mapping pedigree was established by crossing one yellowbeak male and one yellowbeak female with wild type birds from the University of Sheffield’s captive colony [S1]. A panel of 93 birds was typed, including the two YB founders, one of their F0 generation mates, 28 F1 progeny, and 60 F2 birds. Seventeen of the F2 birds expressed the yellowbeak phenotype, and 43 were wildtype birds, which is a ratio consistent with the 3:1 Mendelian expectation when wildtype is dominant to yellowbeak (binomial test, $P = 0.55$). Note that the F2 birds were derived from 8 different F1 crosses between related parents (8 males, 5 females). Birds were humanely killed by cervical dislocation under licence, and the following organs/tissues removed and stored in RNAlater: beak, eye, tarsus, liver, duodenum, leg muscle. Samples from wild birds for genotyping were obtained from the population at Fowlers Gap, New South Wales, Australia.

HPLC

Carotenoid extraction and HPLC analysis of beak tissue samples (upper mandible) from yellowbeak and wildtype male zebra finches largely followed previously published methods for feathers [S3], with the addition of saponification (to hydrolyse carotenoid esters) following Toomey and McGraw [S4]. After tissue homogenization in 2 ml methanol for 15 min at 27Hz, in a Retsch MM2000 ball mill (Hann, Germany), the tissue remains were filtered off with a 0.2-µm syringe filter (GHP Acrodisc), followed by evaporation to dryness under nitrogen in a speedvac (Savant DNA120, Holbrook, USA). The residue was resuspended in 1 ml 0.02 M KOH, incubated in darkness for 6 h at room temperature, and mixed with 0.5 ml saturated NaCl in deionized water. From this solution, free carotenoids were extracted by adding 200 µl ethanol and 200 µl hexane, vortexing, centrifugation for 1 min at 9000 xg, and transferring the hexane supernatant to a new tube. After evaporation and resuspension in 100 µl mobile phase (acetonitrile:methanol 90:10), the samples were immediately analysed isocratically by injecting 50 µl into a Kinetex C18 2.6µ column (Phenomenex, Torrance, USA), on a ThermoFinnigan HPLC system (San Jose, USA) with PS4000 ternary pump, AS3000 autosampler, and UV6000 UV/VIS diode-array detector. Column temperature was maintained at 30°C and the flow-rate at 1.0 ml/min.

For all samples, 2D and 3D chromatograms were obtained, inspected and analysed with ChromQuest 4.0 software (ThermoFinnigan, San Jose, USA). Peaks were identified by comparisons to relative retention times in published analyses of carotenoids in zebra finch
beak and other tissues [18], as well as spectral absorptance shapes obtained from standard runs and earlier analyses on this system (see [S3] for details).

**Genome-wide, low resolution mapping of yellowbeak**

An ideal set of SNPs was selected from the first-generation linkage map [S4]. SNPs were chosen if they had known chromosomal position (not on Chr?_random), and if during construction of the first generation linkage map they had a call frequency >0.6 and a minor allele frequency >0.1. An additional 66 SNPs identified on dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/index.html) were added. These SNPs were chosen to provide greater coverage on the microchromosomes (>5 per microchromosome) or were close to candidate pigmentation genes (26 SNPs). A final panel of 384 SNPs was screened on an Illumina BeadXpress using Goldengate VeraCode technology for the 93-bird mapping panel.

A linkage map was constructed using CRI-MAP (v2.4). We assumed that the zebra finch genome assembly reported the SNPs in the correct order and calculated the recombination distances between SNPs. SNPs that caused map inflation were removed; the total number of SNPs in the linkage map was 352. To determine the position of the loci controlling the recessive yellowbeak mutant, we calculated the two-point linkage between the trait and the 352 retained markers using CRI-MAP. We coded the recessive phenotype as a locus: wild type parents (founders) coded as 1/1, mutant parents as 2/2, all F1 individuals as 1/2, and F2 as 2/2 if they expressed the phenotype and 1/0 if they did not (where 0 represents an uncertain allele). The two-point linkage approach is useful for estimating a coarse chromosomal position. Other multipoint estimates of linkage were not used because the power of the study was limited by the modest number of recombination events that had occurred in the three-generation pedigree, and because two-point analysis revealed a locus perfectly co-segregating with yellowbeak, indicating that multipoint mapping could not improve resolution.

**Fine mapping of yellowbeak**

To fine map the yellowbeak locus we typed 48 SNPs located between 15.02 and 27.28 Mbp on chromosome 8 (see Table S2 for locus information) across the 93-bird mapping panel. Two additional birds of interest were typed; one was a yellowbeak bird that had appeared in our aviary but was not a descendant of the F0 yellowbeak founders of our mapping cross. Genotyping was performed by KASP-genotyping on an LGC SNPline system (http://www.lgcgroup.com/products/genotyping-instruments/snpline/) at the NERC Biomolecular Analysis Facility at the University of Sheffield (NBAF–S). Genotype calling was
performed with the Kraken software (http://www.lgcgroup.com/products/genotyping-software/kraken). Loci were removed prior to downstream analyses if they had a call rate of < 0.70, were out of Hardy-Weinberg Equilibrium at $P < 0.001$, or were considered difficult to score with the Kraken software due to poorly defined genotype clusters. Eight individuals were removed because they had call rates < 0.70. A further individual showed multiple Mendelian inconsistencies with its offspring, and was removed because we suspected a tube-labelling error. The final dataset consisted of 84 birds typed at 39 SNPs.

We tested for linkage between each of the 39 SNPs and the *yellowbeak* locus using CRI-MAP 2.507 (available at http://www.animalgenome.org/tools/share/crimap/), and exactly the same two-point mapping procedure as outlined in the section on Low Resolution Mapping. During the course of our study, an additional bird with a yellow beak appeared in our population. This bird was not a descendant of either of the F0 *yellowbeak* founders (although it presumably shared some recent ancestry with one or both of them). The new bird shared the same homozygous genotypes as all of the *yellowbeak* F0 and F2 birds at a run of 6 consecutive SNPs between 24.01Mbp and 25.68Mbp, which is again consistent with *yellowbeak* being located in the *CYP2J2*-like cluster.

**CYP2J2 cluster gene arrangement**

Long-range PCR was conducted using Extensor kits (Thermo Scientific) under standard conditions, with extension times of 8–10 min. Illumina MiSeq sequencing of PCR amplicons was performed to >1,000-fold coverage, and *de novo* assembly conducted using Seqman NGen (Linux) v.12 (DNASTAR).

The current arrangement of *CYP2J2*-like loci in the zebra finch genome is shown in Figure S1B. There are two loci anchored on chromosome 8 (*LOC24206, LOC24220*), near to *HOOK1*, and two loci on an unanchored chromosome 8 contig (*LOC30349, LOC17987*). None of these loci have official gene names in zebra finch (nor do their orthologues in chicken). They are all related to human *CYP2J2* and we refer to them collectively as *CYP2J2*-like loci. For individual loci we use nomenclature based on recent studies of CYPs in birds ([29, 30] in main reference list). *LOC24220* encodes an intact cytochrome P450 and is divergent from the other three loci, and has been named *CYP2J40*, related to chicken *CYP2J20*, *CYP2J21* and *CYP2J22*. The other three loci are closely related, but only *LOC17987* appears to potentially encode an intact cytochrome P450 and this is the only one of the three that is in Ensembl (ENSTGUG000016903). *LOC30349* has multiple internal duplications, 12 coding exons (CYPs usually have 9) and a premature stop codon. *LOC24206* has a large deletion compared to other CYPs and has recently been discontinued.
We used long-range PCR using primers to unique flanking sequences to demonstrate that some sequences present in the currently unanchored chromosome 8 contig are actually part of the CYP2J2-like cluster on chromosome 8. We sequenced PCR amplicons (~25kbp and ~13kbp) using the same sets of primers shown in Figure S1C using Illumina MiSeq in two wildtype birds. Key evidence is that a PCR amplicon spanning from the 5’ upstream region of the first CYP2J2-like locus on chromosome 8 through to the first exon of CYP2J40 (primers CJ2up9F, CJ1ex1R2) contains two intact CYP2J2-like loci, with a unique sequence from the unanchored contig present just 3’ to the coding sequence of the first locus (Figure S1B). Extensive further long-range PCR experiments failed to uncover any other CYP2J2-like copies, and all cDNA sequences from different tissues matched one of the three copies. Hence we conclude that there are three CYP2J2-like loci at this position on chromosome 8. We refer to the first two copies as CYP2J19A and CYP2J19B since they are paralogues that are related to chicken CYP2J19.

Although CYP2J19A and CYP2J19B are closely related to sequence from LOC24206, LOC17897 and LOC30349, they are distinct from all of them, and it is possible that all of these currently annotated loci result from misassembly.

**CYP2J cluster deletion genotyping**

Long-range PCR using primers CJ2up9F and CJ1ex1R2 revealed a ~13kbp deletion in yellowbeak birds (Figure S1D). We designed a long-range PCR assay to type the ~13kbp deletion in the yellowbeak mapping family. We were unable to obtain 25kbp amplicons (as in Figure S1C) from many samples, presumably because of DNA quality, but instead designed an assay based on multiple 13-kbp amplicons (Figure S4). In this assay, one primer combination (A) gives a 13-kbp amplicon from the yellowbeak allele (and potentially a 25kbp amplicon in wildtype), while two primer combinations (B,C) give 13kbp amplicons from the wildtype allele, and are null for the yellowbeak allele. Hence both homozygotes and heterozygotes could be scored with this assay. We sequenced the 13kbp amplicon in yellowbeak birds using MiSeq.

**Gene expression**

Shavings of the beak sheath, and dissected retinas, were used for RNA extraction. Extraction of total RNA, first strand synthesis with N₆ primers, qualitative RT-PCR and Sanger sequencing of RT-PCR products on both strands were carried out under standard conditions. Quantitative real-time RT-PCR was carried out using Quantitech SYBRGreen kits (Qiagen) in an MJ Opticon2 (Research Engines) thermal cycler. Following previous experience with avian tissues, β-actin was used as a control locus [16]. Quantitation followed our previous methods [16], based on [S5]. The results were taken as averages of three
technical replicates, and PCRs on target and control loci were performed using product from the same cDNA synthesis reaction in the same PCR reaction. Statistical significance was assessed using two-tailed $t$-tests.

**Phylogenetic analysis of CYP2J19 loci**

Genomic sequences were aligned in MEGA6, and phylogenies reconstructed using Maximum Likelihood in MEGA, with a GTR model of sequence evolution. Bootstrap support was estimated using 1000 replicates (Figure S2A).

We considered whether the two divergent allelic lineages in CYP2J19B could actually refer to different loci, but found that this was not the case, since birds homozygous for wildtype alleles could be homozygous or heterozygous for the divergent alleles, and, in particular, birds that were heterozygous for the yellowbeak allele only had one of the divergent wildtype alleles.

**Supplemental References**

S1. Birkhead, T. R., Pellatt, E. J., Matthews, I. M., Roddis, N. J., Hunter, F. M., McPhie, F. and Castillo-Juarez, H. (2006). Genic capture and the genetic basis of sexually selected traits in the zebra finch. Evolution 60, 2389-98.
S2. Andersson S., Prager, M., and Johansson, E. I. A. (2007). Carotenoid content and reflectance of yellow and red nuptial plumages in widowbirds (Euplectes spp.). Funct. Ecol. 21, 272-281.
S3. Toomey, M. B., and McGraw, K. J. (2007). Modified saponification and HPLC methods for analysing carotenoids from the retina of quail: Implications for its use as a nonprimate model species. Invest. Ophth. Vis. Sci. 48, 3976-3982.
S4. Stapley, J., Birkhead, T.R., Burke, T. A., and Slate, J. (2008). A linkage map of the zebra finch Taeniopygia guttata provides new insight into avian genome evolution. Genetics 179, 651-667.
S5. Pfaffl, M. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, e45.