MOLECULAR ECOLOGY

Expression of a carotenoid-modifying gene and evolution of red coloration in weaverbirds (Ploceidae)

Journal:	*Molecular Ecology*
Manuscript ID	MEC-17-1128.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	10-Nov-2017
Complete List of Authors:	Twyman, Hanlu; University of Cambridge, Zoology Prager, Maria; University of Gothenburg, Department of Biological and Environmental Sciences Mundy, Nicholas ; University of Cambridge, Zoology Andersson, Staffan; University of Gothenburg, Department of Biological and Environmental Sciences
Keywords:	Carotenoid metabolism, weaverbirds, CYP2J19, Cytochrome P450

Expression of a carotenoid-modifying gene and evolution of red colouration in weaverbirds (Ploceidae)

[1]Hanlu Twyman (HT), [2]Maria Prager (MP), [1*]Nicholas I. Mundy (NIM) & [2*]Staffan Andersson (SA)

[1]Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK

[2]Department of Biological and Environmental Sciences, University of Gothenburg, Box 463, 40530 Göteborg, Sweden

*Joint senior authors

Corresponding authors:

Staffan Andersson: staffan.andersson@bioenv.gu.se

Nick Mundy: nim21@cam.ac.uk

Abstract

Red carotenoid colours in birds are widely assumed to be sexually selected quality indicators, but this rests on a very incomplete understanding of genetic mechanisms and honesty-mediating costs. Recent progress was made by the implication of the gene *CYP2J19* as an avian carotenoid ketolase, catalysing the synthesis of red C4-ketocarotenoids from yellow dietary precursors, and potentially a major mechanism behind red coloration in birds. Here we investigate the role of *CYP2J19* in the spectacular colour diversification of African weaverbirds (Ploceidae), represented by five genera and 16 species; eight red, seven yellow, and one without carotenoid coloration. All species had a single copy of *CYP2J19*, unlike the duplication found in the zebra finch, with high expression in the retina, confirming its function in coloring red oil droplets. Expression was weak or undetected in skin and follicles of pigment-depositing feather buds, as well as in beaks and tarsi, including those of the red-billed quelea. In contrast, the hepatic (liver) expression of *CYP2J19* was consistently higher (>14 fold) in seven species with C4-ketocarotenoid coloration than in species without (including one red species), an association strongly supported by a phylogenetic comparative analysis. The results suggest a critical role of the candidate ketolase, *CYP2J19,* in the evolution of red C4-ketocarotenoid colour variation in ploceids. Since ancestral state reconstruction suggests that ketocarotenoid coloration has evolved twice in this group (once in *Euplectes* and once in the *Quelea*/*Foudia* clade), we argue that while *CYP2J19* has retained its ancestral role in the retina, it has likely been co-opted for red coloration independently in the two lineages, via increased hepatic expression.

Word count: = 262

Keywords: Carotenoid metabolism, weaverbirds, *CYP2J19,* cytochrome P450

Introduction

Vivid red or yellow colours in birds and other animals are usually carotenoid-based and widely assumed to be sexually or socially selected quality indicators (see e.g., Hill & McGraw 2006a; Svensson & Wong 2011). This assumption rests, however, on a rather incomplete understanding of the underlying physiological and, especially, genetic mechanisms of carotenoid coloration (Toews *et al.* 2017), which are likely to hold the keys to macroevolutionary constraints as well as intraspecific honesty-mediating costs. In a high proportion of cases the red coloration is due to C4-ketocarotenoid pigments, which cannot usually be directly obtained from the diet, but must be synthesized by metabolism of dietary yellow carotenoids (Brush 1990). Ingested yellow carotenoids such as lutein, β-carotene, and zeaxanthin undergo a C4 ketolation reaction, which introduces a double-bonded oxygen (forming a keto-group) at the C4 carbon position of one or both end rings of the carotenoid molecule. This results in a monoketo- or diketo-carotenoid with peak absorptance shifted towards longer wavelengths (and 'redder' hue). From the above precursors, these 'modified red' carotenoids are typically α-doradexanthin, canthaxanthin and astaxanthin (Andersson *et al.* 2007; McGraw 2004; Stradi *et al.* 2001).

The ability to perform carotenoid ketolation is thus likely an important innovation in the evolution and diversification of carotenoid pigments and coloration in vertebrates. In birds, fish and lizards there is much evidence for sexual or social signal selection for red coloration (Hill & McGraw 2006; Ibanez *et al.* 2014; Milinski & Bakker 1990; Svensson & Wong 2011) and even pre-existing receiver biases (Ninnes *et al.* 2017). Despite this, intensely red-colored integument (plumage, beak, skin) has a surprisingly limited and patchy distribution across birds (Aves). Even in clades where red carotenoid coloration is common, e.g. widowbirds and bishops (Prager & Andersson 2010), New World blackbirds (Friedman *et al.* 2014) and cardueline finches (Ligon *et al.* 2016), its absence in several lineages is not associated with any obvious and relevant ecological or behavioral differences from their red-

colored relatives. This suggests that some genetic or physiologically 'hard-wired' constraint is at play, and that C4-ketolation of integumentary carotenoids is likely to be a major hurdle for the evolution of red pigmentation.

A vertebrate C4 ketolase was proposed decades ago (Völker 1962) but its genetic basis has remained unknown. Recently, however, progress was made when the locus *CYP2J19*, of the cytochrome P450 family of monooxygenases, was described as a putative ketolase associated with ketocarotenoid pigmentation in two independent studies of aberrantly colored cage birds: the 'yellowbeak' zebra finch mutant (Mundy *et al.* 2016) and the 'red factor' breed of canary (Lopes *et al.* 2016), in which red coloration was introgressed from the red siskin. The generality of this mechanism, however, has yet to be evaluated since these are the only cases where red integumentary coloration has been linked to *CYP2J19*, in a mutant and a hybrid breeding line, respectively. In addition they differ in a) *CYP2J19* gene copy number (two in zebra finch and one in the red factor canary) and b) tissue expression ('peripherally' in the integument in zebra finch and both 'peripherally and centrally' in the feather follicles and liver in the 'red factor' canary). There are thus many remaining questions concerning the generality, nature and location of the *CYP2J19* mechanism and function in red coloration. More broadly, *CYP2J19* appears to be conserved for retinal red oil droplet pigmentation within the turtles (the only other group of tetrapods to possess red oil droplets apart from the birds), from which it has been recruited for red integumentary coloration independently within certain turtle and avian lineages (Twyman et al. 2016). In particular, it remains unknown whether differential *CYP2J19* expression can account for variation in red coloration across and between avian clades.

Weaverbirds (Aves: Ploceidae, 116 species) are a clade of predominantly African, seed-eating passerines, which are ideal for studying the mechanisms and evolution of carotenoid coloration. Whereas conspicuous yellow plumage colours dominate, especially in the most speciose lineages of 'true weavers' (*Ploceus* spp), red carotenoid coloration occurs

in several genera, and a few lineages lack integumentary carotenoid pigmentation altogether. The underlying mechanisms (e.g. dietary vs metabolically modified pigments) have been established for several species, notably the brilliant yellow or red plumage displays of widowbirds and bishops (*Euplectes*) (Andersson *et al.* 2007; Prager *et al.* 2009). In this clade, red colour hues are agonistic (threat) signals used in male contest competition, and appear to have evolved at least twice from a yellow ancestor (Prager & Andersson 2010) and apparently due to a pre-existing receiver bias for redder (longer wavelength) hues (Ninnes *et al.* 2017). Outside *Euplectes*, weaverbird colour signalling functions are largely unexplored, except in the red-billed quelea (*Quelea quelea*), where the red beak likely is sexually selected (through either female mate choice or male-male contests) whereas the polymorphic red plumage coloration may be involved in individual recognition (Dale 2000).

In most ploceids where it has been analysed, red colour patches contain red C4-ketocarotenoids, primarily α-doradexanthin and canthaxanthin, co-deposited with the dietary yellow precursor pigments (Andersson *et al.* 2007; unpublished results). By comparisons to phylogenetically, socially and ecologically closely related yellow-colored species, this provides an excellent opportunity to test the significance of *CYP2J19* for red carotenoid coloration. Moreover, the fantailed widowbird (*E. axillaris*) has been found to achieve its striking red wing patch coloration without C4-ketocarotenoids (Andersson *et al.* 2007; Prager *et al.* 2009), which provides an additional test of the proposed function (C4-ketolation) of *CYP2J19*.

In this study of 16 red or yellow weaverbird species, we investigate the role of *CYP2J19* in the evolution of carotenoid pigmentation in weaverbirds. First, we establish whether the gene is present in ploceids and, if so, in how many copies. Second, we identify the anatomical site(s) of *CYP2J19* expression in this group. Finally, using a phylogenetic comparative analysis, we test whether *CYP2J19* expression is associated with the occurrence of red C4-ketocarotenoid pigmentation across the ploceids.

Materials and Methods

Samples

Feathers (for HPLC) and tissue samples (for qRT-PCR) from male ploceids in breeding plumage were largely obtained from natural populations in Africa, in addition to a few samples from aviaries in southern Spain and Sweden (Table 1), under all applicable national and international permits. 5-10 feathers were plucked with flat-tipped tweezers and stored in dark envelopes until analysis. Euthanised birds were freshly dissected and tissues placed in RNAlater (Qiagen) or DNA/RNA-shield (Zymo) until DNA/RNA extraction. Follicles for gene expression analysis were sampled from growing, carotenoid-depositing feather buds.

C4-ketocarotenoid pigmentation

The presence of integumentary C4-ketocarotenoid pigments (Figure 2) was established from published HPLC (High Performance Liquid Chromatography) analyses of feathers and beak tissue from six of the included species: *E. ardens*, *E. axillaris*, and *E. macroura* (Andersson *et al.* 2007), *E. afer* and *E. orix* (Prager *et al.* 2009) and *Q. quelea* (Walsh *et al.* 2012). For the remaining species in this study, C4-ketocarotenoid presence or absence was determined from unpublished HPLC analyses performed in conjunction with the above studies, using identical or very similar methods (see Supporting Methods, Supporting Table 1).

CYP2J19 gene copy number determination

Genomic DNA was extracted from liver using QIAamp DNA Mini kits (Qiagen) according to the standard protocol. Long-range PCR was conducted to determine *CYP2J19* gene copy number with a published protocol (Mundy *et al.* 2016) using Extensor kits (Thermo Scientific) under standard conditions, with extension times of 8-10 minutes. The size of long-range amplicons was measured using Quick-Load 1kb Extend DNA Ladder (BioLabs) on 0.6% agarose gels. Illumina MiSeq sequencing of PCR amplicons was performed to >1,000-fold

coverage at the University of Sheffield and *de novo* assembly was conducted using Seqman NGen (Linux) v.12 (DNASTAR) for *Q. quelea* and *E. orix*. Genomic sequences have been deposited in GenBank (Accession MG255072-3).

CYP2J19 expression

Total RNA was extracted from all tissue samples using RNeasy Mini kits (Qiagen). Dissected tissues were manually homogenized using an Eppendorf homogenizer prior to addition of Buffer RLT. The lysate was centrifuged for 2 minutes at 13,000rpm in QIAshredder spin columns before proceeding with subsequent full speed centrifugation step for 3 minutes. DNase digestion was performed using Qiagen RNase-Free DNase Set.

First strand synthesis was performed with 10µl total RNA and N6 primer (0.5µM) using SuperScriptII RT (Life technology Invitrogen) according to the manufacturer's instructions. RT-PCR reactions contained 1 x NH4 Buffer, MgCl2 (1.5mM), each dNTP (2.5mM), each primer (0.4µM), BioTaq DNA polymerase (Bioline) (0.5U) and cDNA (~50ng). Reactions were run in a G-Storm GS1 Thermal Cycler (Life Science Research) under the following conditions: 2 minutes at 94°C followed by 35–40 cycles of heating for 30 seconds at 94°C, 45 seconds at 60°C and 90 seconds at 72°C with a final extension of 5 minutes at 72°C. The amplified full length fragment was purified using ExoSap-IT (Affymetrix) and sequenced on both strands via Sanger sequencing. cDNA sequences have been deposited in Genbank (Accessions MG255074-86).

Quantitative real-time RT-PCR was carried out in an MJ Opticon2 (Research Engines) thermal cycler using the Quantitech SYBRGreen kit (Qiagen), using three reference genes (*β-Actin, GAPDH* and *HPRT1*) and three technical replicates for each condition. Tissue specific expression differences between four tissue types of *Q. quelea* were assessed using Analysis of Variance under the 'car' package in R version 3.3.2 (R Core Team 2016), and the Box-Cox power transformation for normality was applied, with lambda fixed at 0.1.

The Shapiro-Wilk test for normality of residuals ($p>0.8$), Bartlett's test for equality of variance ($p>0.1$) and Runs test for independence of residuals ($p>0.06$) were upheld.

Normalisation following Pfaffl (2001) was performed using *β-Actin* in the first round of analyses and, in the final analyses, using the geometric mean of the three reference loci. The geNorm application for the evaluation of expression stability in the control genes was applied to assess the suitability of the reference loci (Vandesompele *et al.* 2002). The M values (denoted as the average pairwise variation of a control gene with all other control genes) for *β-Actin, GAPDH* and *HPRT1* were 1.185, 1.142, and 1.107 respectively, indicating suitability for their use.

Association between CYP2J19 and red ketocarotenoid pigmentation

To account for phylogenetic non-independence between species, the association between hepatic *CYP2J19* expression and red ketocarotenoid pigmentation was assessed using the discrete Markov chain Monte Carlo (MCMC) method in BayesTraits V2 (Pagel 1994; Pagel *et al.* 2004, www.evolution.rdg.ac.uk), while sampling from 10,000 Ploceidae trees downloaded from BirdTree.org (Jetz *et al.* 2012, 'Ericson All Species' source). A phylogeny of Ploceidae was recently published (De Silva et al. 2017), but contained several identification errors and interspecific sequence concatenations (Prager 2017), some of which involve taxa included here. The BirdTree phylogeny is, in contrast, congruent with a previous phylogeny of the genus *Euplectes*, including a few *Quelea*, *Foudia* and *Ploceus* taxa (Prager et al 2008).

Log transformed normalised values of liver *CYP2J19* expression were first discretized using k-means clustering in R version 3.3.1 (R Core Team 2016) with two cluster centres ('high' and 'low'), excluding *Euplectes aureus*, *E. axillaris* and *E. macrourus* where expression was undetectable after 50 PCR cycles (Figure 3). Based also on results from the *β-Actin*-normalised analyses (Supporting Figure 1), the latter were manually scored as 'low'.

In BayesTraits V2, an 'independence model' estimating four separate evolutionary rates (gain and loss for each trait) was compared to a 'dependence model' allowing for a

maximum of eight separate rates (Figure 1). Assuming, however, that red ketocarotenoid pigmentation is contingent on high hepatic *CYP2J19* expression, the rates of all (four) changes involving a state of ketocarotenoid presence at low CYP expression were set to zero in the final dependence model. Marginal likelihoods of alternative models were approximated using harmonic means of log likelihoods from 1 million generations of discrete Markov Chain Monte Carlo (MCMC) runs, discarding the first 100,000 generations as burn-in, and compared using Bayes Factors (Kass & Raftery 1995). As MCMC methods can be sensitive to choice of priors, we tested five different prior distributions on transition rates: 1) the default setting of uniform(0, 100), 2) a conservative 'empirical' prior of uniform(0,1) covering the range of maximum likelihood rates estimated for each tree under the independence model, and 3-5) exponential distributions centred at 0.1, 1 and 10, respectively. Acceptance rates for rate change parameters were confirmed to range within 20-40% to ensure proper mixing of MCMC chains in each model.

Results

Integumentary carotenoid pigmentation

Based on published or hitherto unpublished HPLC analyses of carotenoids in colored feathers or beak tissue of all 16 ploceids (Supporting Table 3), each species was categorized as either 'KC present' (> 0% C4-ketocarotenoids) or 'KC absent' (no C4-ketocarotenoids detected). Whereas only presence/absence of C4-ketocarotenoids ('modified red') was analyzed in relation to *CYP2J19* expression (below), it can be noted that in all seven 'KC present' species, it is the same set of five C4-ketocarotenoids (α-doradexanthin, β-doradexanthin, adonirubin, canthaxanthin, and astaxanthin) but in variable absolute and relative amounts. Only one species, *Quelea quelea*, seems to lack one of the C4-ketocarotenoids, β-doradexanthin. Also noteworthy, as pointed out in Andersson et al (2007), the one species with red coloration without any C4-ketocarotenoids, *Euplectes axillaris,* has 2-3 times as high total concentration of carotenoids and is also the only species with the 'modified yellow' carotenoid anhydrolutein. See Supporting Table 1 for further details.

CYP2J19 presence, copy number and variation

Tissue samples from 16 species of weaverbirds (seven *Euplectes*, five *Ploceus*, two *Quelea*, one *Foudia*, one *Philetairus*) were analysed (see Table 1). Based on a long-range PCR assay on genomic DNA, all 16 species were found to have a single *CYP2J19* gene copy of ~10-15kb, which was confirmed by Illumina Miseq sequencing in two species (*E. orix* and *Q. quelea*). Given the possibility of differential expression of copies in different tissues (as in the zebra finch), we further confirmed a single copy in ploceids by showing that full length sequences of *CYP2J19* cDNA from different tissues (retina and liver) of the same individual were identical, in three species (*E. ardens, F. madagascariensis,* and *Q. quelea*).

Full length *CYP2J19* cDNA sequences revealed that there were no amino acid substitutions shared among species with C4-ketocarotenoid coloration that were not present among species without C4-ketocarotenoids.

Patterns of CYP2J19 expression

Analysis of expression using qualitative RT-PCR showed strong *CYP2J19* expression in the retinas of all species examined (N = 10 species), variable expression in the liver, and weak or undetectable expression in all peripheral tissues (skin and feather follicles (N = 5 species), beak (N = 6), and tarsus (N = 2), Table 2). Using *Q. quelea* as an example since it is the only sampled species with red bare body parts (beak and tarsus), significantly higher expression of *CYP2J19* was found in the liver and retina compared to the beak and tarsus (Figure 1, Table 3).

Initial qRT-PCR quantification of hepatic expression of *CYP2J19* using a single control locus (*β-actin*) and samples of 1-3 breeding males across the 16 species showed high levels of *CYP2J19* in four members of the *Euplectes* clade (*E. orix*, *E. hordeaceus*, *E. nigroventris*, *E. ardens*), two queleas and a fody, with levels > 100-fold greater than all other species (Supporting Figure 1). We confirmed these findings by performing qRT-PCRs using three control loci on a randomly chosen subset of samples (one per species) (Figure 3). These gave similar results, with the same seven species showing high (0.1 – 8.6) levels of hepatic *CYP2J19* compared to the remaining species (<0.007) (>14 fold difference).

Association between CYP2J19 and red ketocarotenoid pigmentation

There is a perfect association between high hepatic *CYP2J19* expression and the presence of red C4-ketocarotenoids: breeding males of the seven species with high liver *CYP2J19* all have red coloration due to red C4-ketocarotenoid pigments (Figure 3). In contrast, the nine species without C4-ketocarotenoids (eight of which have yellow carotenoids, one with no

carotenoids) all have low hepatic *CYP2J19* expression, and this includes *E. axillaris*, the only species sampled here that produces a red colour hue based on 'yellow' carotenoids alone, i.e. without using C4-ketocarotenoids (Andersson *et al.* 2007).

Phylogenetic comparative tests of correlated evolution between hepatic *CYP2J19* expression and red C4-ketocarotenoid pigmentation were performed in BayesTraits V2. Estimated marginal likelihoods, based on five different prior assumptions of transition rates (Table 4), consistently support a ‘dependence model’, where the evolution of red C4-ketocarotenoid pigmentation is contingent on high hepatic expression of *CYP2J19*, over an ‘independence model’, where the rate of change in one trait is unaffected by the state of the other trait (Figure 2). Even with the most conservative priors (i.e. in favour of the 'independence model'), Bayes Factor test statistics (calculated as 2*[lnL(dependent model) - lnL(independent model)] exceeded 10 which is usually interpreted as very strong support for an association (Kass & Raftery 1995).

Discussion

Our results suggest that hepatic expression of *CYP2J19,* a candidate carotenoid ketolase, constitutes a principal mechanism and evolutionary innovation behind red carotenoid coloration in weaverbirds (Ploceidae). Since the interspecific association between high *CYP2J19* expression and presence of red C4-ketocarotenoid pigments could be due to phylogenetic non-independence (shared ancestry), the relationship was tested in a Bayesian phylogenetic comparative analysis and found to be very strong. Our results strengthen *CYP2J19* as the prime candidate for the long-sought avian C4-ketolase.

We have furthermore established that weaverbirds consistently seem to have a single copy of *CYP2J19*. In contrast, the zebra finch, an estrildid finch belonging to the nearest outgroup clade to Ploceidae (De Silva *et al.* 2017), has two copies, *CYP2J19A* and *CYP2J19B*, seemingly specialised for retinal oil droplet pigmentation and integumentary coloration, respectively (Mundy *et al.* 2016). It therefore appears that the estrildid *CYP2J19* duplication occurred after the split between ploceids and estrildids. More broadly, a single copy of *CYP2J19* was reported also in the red factor canary (Lopes *et al.* 2016), as well as in chicken and ostrich (Twyman *et al.* 2016) and GenBank searches reveal only a single copy in the vast majority of available avian genomes (Emerling 2017, Twyman et al. in review), which means that a single *CYP2J19* copy probably is the “normal” situation for birds.

The tissue-specific expression data for *CYP2J19* strongly implicate the liver as the main site of carotenoid ketolation. As earlier suggested by high plasma concentration of red ketocarotenoids (Prager *et al.* 2009; unpublished results), ploceids thus seem to be “central” ketoconverters. Notably in this context, the hepatic *CYP2J19* expression was very low in *E. axillaris*, a species with red carotenoid coloration that does not involve C4-ketocarotenoids (Andersson *et al.* 2007). Apart from implying an intriguing alternative "redness mechanism" (possibly related to the in birds unusual presence of 'anhydrolutein'; see McGraw et al. 2002; Andersson *et al.* 2007), it supports in this context a causative and direct link (i.e. not an

297 indirect association via colour) between hepatic *CYP2J19* expression and C4-
298 ketocarotenoids in the plumage.

299 In contrast to the liver, *CYP2J19* expression was very low or undetectable in
300 peripheral tissues (skin, feather follicles, beak, tarsus), including the red and ketocarotenoid-
301 pigmented beak and legs of the red-billed quelea. Nevertheless, more extensive and careful
302 sampling, covering a broader range of feather growth stages, will be required to rule out the
303 possibility of a 'peripheral' (integumentary) role of *CYP2J19* in feather follicles, as implicated
304 in the red factory canary (Lopes *et al.* 2016).

305 There was substantial variation in hepatic *CYP2J19* expression overall, not least
306 among the ketocarotenoid-colored species, where by far the highest expression was found in
307 the red-billed quelea. Since this is the only of our study species that has a red-colored beak
308 (and tarsi), we speculate that, compared to plumage this continually renewing tissue may
309 require a more constant and larger supply of ketocarotenoids to maintain its red colour. Most
310 of the variability of *CYP2J19* expression among red species, however, had no such obvious
311 association with phenotype, and probably relates to timing of sampling in relation to the pre-
312 nuptial plumage moult, or to some other genetic, social or environmental factor. For example,
313 given that cytochrome P450 enzymes often are regulated by substrate availability (Zanger &
314 Scwab 2013), *CYP2J19* expression is likely affected by both amount and composition of
315 carotenoids in the diet. Further studies of inter- as well as intraspecific variation in *CYP2J19*
316 expression, with carefully controlled and standardized sampling, are needed to explore if
317 some of this variation is biologically meaningful, for example by suggesting physiological
318 costs or trade-off's with detoxification (see Mundy et al 2016) that may mediate honest
319 signalling.

320 Historically there has been considerable debate over the anatomical site of ketolation
321 (McGraw 2004; del Val *et al.* 2009) and even with a few examples it is now apparent that
322 there is substantial variation in the strategy employed by different passerine species. The

contrast between the red-billed quelea and zebra finch, which both have red beak and tarsus, is particularly striking: the former has high *CYP2J19* expression in liver and low/absent expression in beak and tarsus while the zebra finch shows the opposite pattern.

Unlike the situation with two copies of *CYP2J19* in the zebra finch (Mundy et al. 2016), an estrildid that uses peripheral ketoconversion to color its bill and tarsi, the 16 ploceids in this study all have a single copy of *CYP2J19* and central (liver) ketoconversion, supplying either plumage or, in the red-billed quelea, beak and bare part coloration. The red factor canary (a hybrid fringillid), likewise has a single *CYP2J19* copy with both central and peripheral expression (Lopes et al 2016), although this needs to be confirmed in a natural fringillid species. Broader sampling and further study of interspecific variation in *CYP2J19* copy number and site(s) of action may yield interesting evolutionary implications as regard micro- and macroevolutionary constraints on colour and pattern diversification.

Based on a previous ancestral character state reconstruction (Prager & Andersson 2010), the two clades with high hepatic *CYP2J19* expression (*Foudia/Quelea* and *E. ardens/E. hordeaceus/E. nigroventris/E. orix*) likely acquired red coloration convergently. We therefore hypothesise that convergent evolution of red coloration arose in these two clades via increases in hepatic *CYP2J19* expression. Due to the highly conserved function of *CYP2J19* for red retinal oil droplets, this would have required specific acquisition of high liver expression of *CYP2J19* while maintaining high retinal expression, which may have occurred via evolution of *cis*-regulation of *CYP2J19* and/or *trans*-regulating factors. Such co-option of a '4-oxygenase' (i.e. the ketolase), mediated by *cis*-regulatory elements, was also suggested to explain the evolution of C4-ketocarotenoid pigmentation in *Colaptes* woodpeckers (Hudon *et al.* 2015). Given the relatively rare but phylogenetically widespread occurrence of red carotenoid coloration, the co-option of *CYP2J19* seems to have occurred several times independently in birds and also in turtles (Twyman *et al.* 2016), but a scenario with early gains and multiple subsequent losses may well emerge as further lineages are investigated.

It is also important to note that we have only considered a single aspect of carotenoid coloration, the presence of integumentary C4-ketocarotenoids; several other mechanisms, for e.g. uptake, metabolism, transport, and deposition, may also be key factors behind interspecific colour variation.

Cis-regulatory evolution is often regarded as a major motor behind adaptive change (Stern & Orgogozo 2008). Factors contributing to this include the high evolvability of transcription factor binding sites and the avoidance of negative pleiotropic effects (Wittkopp & Kalay, 2012; Aguilar-Rodriguez *et al.* 2017). The evolution of the role of *CYP2J19* in red coloration fits this paradigm well, since it appears to primarily involve a change in tissue-specific gene expression. On the other hand, the likely high evolvability of *CYP2J19* expression rhymes less well with the relatively rare and patchy distribution of red carotenoid coloration in birds, particularly striking in the weaverbirds where there seems to be universal selection and convergent evolution of red color signals (Ninnes *et al.* 2015; Ninnes & Andersson 2014; Prager & Andersson 2010). This may indicate that other locus-specific factors contribute to the constraint, which may include coordination of expression in relation to age, sex, body condition and season, potentially requiring the evolution of multiple cis-regulatory modules for *CYP2J19*. Moreover, red ketocarotenoid based coloration also has a sparse distribution amongst turtles (the only non-avian group shown to possess *CYP2J19*), and whereas less is known about selection for red coloration in this group, locus-specific genetic constraints may also explain some of the patterns of interspecific color variation in the turtles. Hence, elucidating and disentangling potential constraints on the evolution of carotenoid coloration in animals will require detailed investigation of the genetic and environmental causes and consequences of co-opting the *CYP2J19* for integumentary pigmentation.

Rapid progress has recently been made in documenting the genetic basis of convergent evolution of naturally selected traits (Stern 2013). For example, in birds, evolution

of melanin-based coloration in birds is frequently due to two loci, *MC1R* and *ASIP* (Mundy 2005; Toews *et al.* 2017; Uy *et al.* 2016). Here we have uncovered one of the first examples in vertebrates where a locus is involved in convergent evolution of a sexually selected trait. Future work on *CYP2J19* promises many novel insights into both function and evolution of carotenoid coloration in birds, as well as general questions regarding diversification due to differential selection or differential constraints.

Acknowledgements

We are grateful to Nancy Bunbury, Jerome Fuchs, Calum Ninnes, Claire Spottiswoode, José Tella and Christer Wiklund for assistance with obtaining tissue samples, to ECOFAC in Sao Tomé and Principe, Ezemwelo KZN Wildlife and Seychelles Bureau of Standards for collection permits, Maria von Post for help with HPLC analyses and Kang-Wook Kim for assembling MiSeq data. We thank the BBSRC Doctoral Training Partnerships (to HT), Murray Edwards College (to NM) and Swedish Science Council (to SA) for funding.

References

Aguilar-Rodríguez, J., Payne, J. L., & Wagner, A. (2017). A thousand empirical adaptive landscapes and their navigability. *Nature Ecology & Evolution,* **1**, 0045.

Andersson S, Prager M, Johansson EIA (2007) Carotenoid content and reflectance of yellow and red nuptial plumages in widowbirds (Euplectes spp.). *Functional Ecology* **21**, 272-281.

Brush AH (1990) Metabolism of carotenoid pigments in birds. *FASEB Journal* **4**, 2969-2977.

Dale J (2000) Ornamental plumage does not signal male quality in red-billed queleas. *Philosophical Transactions of the Royal Society B: Biological Sciences* **267**, 2143-2149.

Friedman NR, McGraw KJ, Omland KE (2014) Evolution of carotenoid pigmentation in caciques and meadowlarks (Icteridae): repeated gains of red plumage coloration by carotenoid C4-oxygenation. *Evolution* **68**, 791-801.

Hill GE, McGraw KJ (2006) Bird Coloration, vol 2. Function and Evolution. Harvard University Press, Cambridge, USA.

Hudon J, Wiebe KL, Pini E, Stradi R (2015) Plumage pigment differences underlying the yellow-red differentiation in the Northern Flicker (Colaptes auratus). *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology* **183**, 1-10.

Ibanez A, Polo-Cavia N, Lopez P, Martin J (2014) Honest sexual signaling in turtles: experimental evidence of a trade-off between immune response and coloration in red-eared sliders Trachemys scripta elegans. *Naturwissenschaften* **101**, 803-811.

Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO (2012) The global diversity of birds in space and time. *Nature* **491**, 444-448.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association* **90**, 773-795.

414 Ligon RA, Simpson RK, Mason NA, Hill GE, McGraw KJ (2016) Evolutionary innovation and
415 diversification of carotenoid-based pigmentation in finches. *Evolution* **70**, 2839-2852.
416 Lopes RJ, Johnson JD, Toomey MB*, et al.* (2016) Genetic Basis for Red Coloration in Birds.
417 *Curr. Biol.* **26**, 1427-1434.
418 McGraw KJ (2004) Colorful songbirds metabolize carotenoids at the integument. *Journal of*
419 *Avian Biology* **35**, 471-476.
420 McGraw KJ, Adkins-Regan E, Parker RS (2002) Anhydrolutein in the zebra finch: a new,
421 metabolically-derived carotenoid in birds. *Comparative Biochemistry and Physiology*
422 *B-Biochemistry & Molecular Biology* **132**, 811-818.
423 Milinski M, Bakker CM (1990) Female sticklebacks use male coloration in mate choice and
424 hence avoid parasitized males. *Nature* **344**, 330-333.
425 Mundy NI (2005) A window on the genetics of evolution: MC1R and plumage colouration in
426 birds. *Proc. R. Soc. B* **272**, 1633-1640.
427 Mundy NI, Stapley J, Bennison C*, et al.* (2016) Red Carotenoid Coloration in the Zebra Finch
428 Is Controlled by a Cytochrome P450 Gene Cluster. *Curr. Biol.* **26**, 1435-1440.
429 Ninnes CE, Adrion M, Edelaar P, Tella JL, Andersson S (2015) A receiver bias for red
430 predates the convergent evolution of red color in widowbirds and bishops. *Behavioral*
431 *Ecology* **26**, 1212-1218.
432 Ninnes CE, Andersson S (2014) Male receiver bias for red agonistic signalling in a yellow-
433 signalling widowbird: a field experiment. *Proceedings of the Royal Society B:*
434 *Biological Sciences* **281**.
435 Ninnes CE, Webb SL, Andersson S (2017) Are red bishops red enough? On the persistence
436 of a generalized receiver bias in Euplectes. *Behav. Ecol.* **28**, 117-122.
437 Pagel M (1994) Detecting Correlated Evolution on Phylogenies - a General-Method for the
438 Comparative-Analysis of Discrete Characters. *Proceedings of the Royal Society of*
439 *London Series B-Biological Sciences* **255**, 37-45.

Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology* **53**, 673-684.

Prager M, Andersson S (2010) Convergent evolution of red carotenoid coloration in widowbirds and bishops (Euplectes spp.). *Evolution* **64**, 3609-3619.

Prager M, Johansson EIA, Andersson S (2009) Differential ability of carotenoid C4-oxygenation in yellow and red bishop species (Euplectes spp.). *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology* **154**, 373-380.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Stern DL (2013) The genetic causes of convergent evolution. *Nat. Rev. Genet.* **14**, 751-764.

Stradi R, Pini E, Celentano G (2001) Carotenoids in bird plumage: the complement of red pigments in the plumage of wild and captive bullfinch (Pyrrhula pyrrhula). *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology* **128**, 529-535.

Svensson PA, Wong BBM (2011) Carotenoid-based signals in behavioural ecology: a review. *Behaviour* **148**, 131-189.

Toews DP, Hofmeister NR, Taylor SA (2017) The Evolution and Genetics of Carotenoid Processing in Animals. *Trends Genet* **33**, 171-182.

Twyman H, Valenzuela N, Literman R, Andersson S, Mundy NI (2016) Seeing red to being red: conserved genetic mechanism for red cone oil droplets and co-option for red coloration in birds and turtles. *Proc Biol Sci* **283**.

Uy JAC, Cooper EA, Cutie S*, et al.* (2016) Mutations in different pigmentation genes are associated with parallel melanism in island flycatchers. *Proceedings of the Royal Society B-Biological Sciences* **283**.

Val E, Senar JC, Garrido-Fernández J*, et al.* (2009) The liver but not the skin is the site for conversion of a red carotenoid in a passerine bird. *Naturwissenschaften* **96**, 797-801.

Vandesompele J, De Preter K, Pattyn F*, et al.* (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, research0034.0031-research0034.0011.

Völker O (1962) Experimentelle Untersuchungen zur Frage der Entstehung roter Lipochrome in Vogelfedern: Farbenfütterungsversuche am roten Kanarienvogel. *Journal für Ornithologie* **103**, 276-286.

Walsh N, Dale J, McGraw KJ, Pointer MA, Mundy NI (2012) Candidate genes for carotenoid coloration in vertebrates and their expression profiles in the carotenoid-containing plumage and bill of a wild bird. *Proceedings of the Royal Society B-Biological Sciences* **279**, 58-66.

Zanger UM, Schwab M (2013) Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities and impact of genetic variation. *Pharmacol. Therapeutics* **138**, 103-141.

Data Accessibility

DNA sequences: Genbank accession MG255072-86

Author Contributions

HT designed and carried out molecular laboratory work, analysed the data and helped edit the manuscript; MP analysed HPLC data, performed phylogenetic comparative analysis and edited the manuscript; SA collected samples, designed and carried out HPLC analyses, and together with NIM conceived the study, designed the experiments and drafted the manuscript. All authors gave final approval for publication.

Tables and Figures

Table 1. Samples used for molecular analysis.

Species	Carotenoid-based coloration	Individual	Tissue	Origin	Date	Collector(s)
Philetairus socius	N	1	†‡L	Benfontein Nature Reserve, South Africa	Nov-2004	SA
Ploceus subaureus	Y	1	†‡L	Salima, Malawi	Dec-2004	SA
Ploceus melanocephalus	Y	1	†‡L, R	Sanlúcar la Mayor, Spain	Sep-2012	CN, JT
Ploceus capensis	Y	1	†‡L, R, B	Port Elizabeth, South Africa	Oct-2014	SA
Ploceus velatus	Y	1	†‡L, R, B	Port Elizabeth, South Africa	Oct-2014	SA
	Y	2	L, R	Blouberg Nature Reserve, Limpopo, South Africa	Oct-2013	JF
Foudia madagascariensis	R	1	†‡L	Assumption Island, Seychelles	Apr-2012	NB
	R	2	†L	Assumption Island, Seychelles	Apr-2012	NB
	R	3	†L, R, B	Mahé, Seychelles	Nov-2014	SA, MP
	R	4-5	L, R, B	Mahé, Seychelles	Nov-2014	SA, MP

Quelea erythrops	R	1	††L, S+F	São Tomé, S.T. and Príncipe	Jan-2008	SA, MP, NIM
Quelea quelea	R	1-2	†L, †R, †B, †T	Zambia	Nov-2012	CS
	R	3	††L, †R, †B, †T	Zambia	Nov-2012	CS
Euplectes afer	Y	1	††L, R	Sanlucar la Mayor, Spain	Sep-2012	CN, JT
	Y	2	S+F	Unknown (commercially obtained)	May-2006	SA
	Y	3	S+F	Unknown (commercially obtained)	Feb-2006	SA
Euplectes aureus	Y	1	††L, S+F	São Tomé, S.T. and Príncipe	Nov-2007	SA, MP
	Y	2-3	F	São Tomé, S.T. and Príncipe	Nov-2007	SA, MP
	Y	4	F	São Tomé, S.T. and Príncipe	Jan-2008	SA, MP, NIM
Euplectes axillaris	R	1	††L, R	Pietermaritzburg, South Africa	Oct-2013	SA
	R	2	L, R	Pietermaritzburg, South Africa	Oct-2013	SA
Euplectes macroura	Y	1	†L	Buea, Cameroon	Jul-2012	CW
	Y	2	††L, R, B	Choma, Zambia	Nov-2012	CS
	Y	3	†L, R, B	Choma, Zambia	Nov-2012	CS
Euplectes ardens	R	1	†L	Iringa, Tanzania	Feb-2011	SA, MP
	R	2	†L	KwaZulu-Natal, South Africa	May-2006	SA
	R	3	††L, R	Cedara, South Africa	Oct-2013	SA
	R	4	R	Cedara, South Africa	Oct-2013	SA
	R	5-6	F	Iringa, Tanzania	Feb-2011	SA, MP
Euplectes hordeaceus	R	1	‡L	São Tomé, S.T. and Príncipe	Nov-2007	SA, MP
	R	2	†L, F	São Tomé, S.T. and Príncipe	Nov-2007	SA, MP

	R	3-4	S+F	São Tomé, S.T. and Príncipe	Nov-2007	SA, MP
Euplectes nigroventris	R	1	††‡L, S+F	Unknown (commercially obtained)	Apr-2006	SA
Euplectes orix	R	1	R	Cedara, South Africa	Oct-2013	SA
	R	2-3	†L	Cedara, South Africa	Oct-2013	SA
	R	4	††‡L, B, T	Cedara, South Africa	Oct-2013	SA

Carotenoid-based coloration (irrespective of C4-keto-carotenoid presence): Y = yellow, R = red, N = absent
Sampled tissues: B = beak, F = feather follicle, L = liver, R = retina, S+F = skin + feather follicle, T = tarsus
‡Samples used for qRT-PCR normalised against 3 reference loci (*β-Actin, GAPDH* and *HPRT1)*
†Samples used for qRT-PCR normalised against 1 reference locus (*β-Actin)* – see Supporting Information for results
Tissue collector(s): SA = Staffan Andersson, NB = Nancy Bunbury, JF = Jerome Fuchs, NIM = Nicholas Mundy, CN = Calum Ninnes, MP = Maria Prager, CS = Claire Spottiswoode, JT = José Tella, CW = Christer Wiklund

Table 2. Qualitative analysis of *CYP2J19* expression in the retina, beak, tarsus, skin and feather follicles of 16 ploceid species.

SPECIES	TISSUES			
	Retina	Beak	Tarsus	Skin and feather follicle
Ploceus melanocephalus	●			
Ploceus capensis	●	○		
Ploceus velatus	●	○		
Foudia madagascariensis	●	-		
Quelea erythrops				○
Quelea quelea	●	○	○	
Euplectes afer	●			○
Euplectes aureus				○
Euplectes axillaris	●			
Euplectes macroura	●	-		
Euplectes ardens	●			○
Euplectes hordeaceus				○
Euplectes nigroventris				○
Euplectes orix	●	○	○	

Strong (●), weak (○) and undetectable (-) expression levels are shown. Gaps in the table were not determined.

Table 3. Tukey's pairwise tests of *CYP2J19* expression for four *Q. quelea* tissues

Pairwise tissue comparisons	p adj
Liver > Beak	< 0.001 ***
Retina > Beak	< 0.001 ***
Tarsus ~ Beak	0.167
Liver > Retina	0.003 **
Liver > Tarsus	< 0.001 ***
Retina > Tarsus	< 0.001 ***

Table 4. Estimated marginal log-likelihoods (lnL) of 'dependency' (dep) versus 'independency' (indep) models, and the Bayes Factor test statistic (2lnBF), given different prior assumptions of evolutionary transition rates.

Rate distribution prior	lnL(dep)	lnL(indep)	2lnBF
Uniform (0, 100)	-11.6	-21.8	20.5
Uniform (0, 1)	-12.0	-19.4	14.8
Exponential (1/λ=0.1)	-11.2	-16.6	10.8
Exponential (1/λ=1)	-11.2	-20.2	18.1
Exponential (1/λ=10)	-10.9	-22.0	22.1

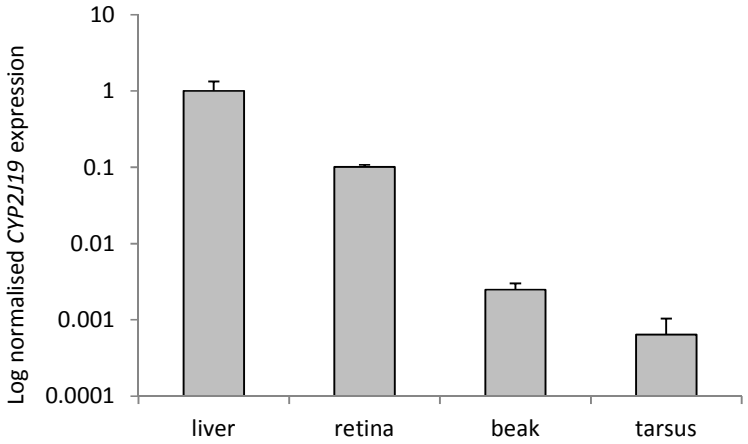
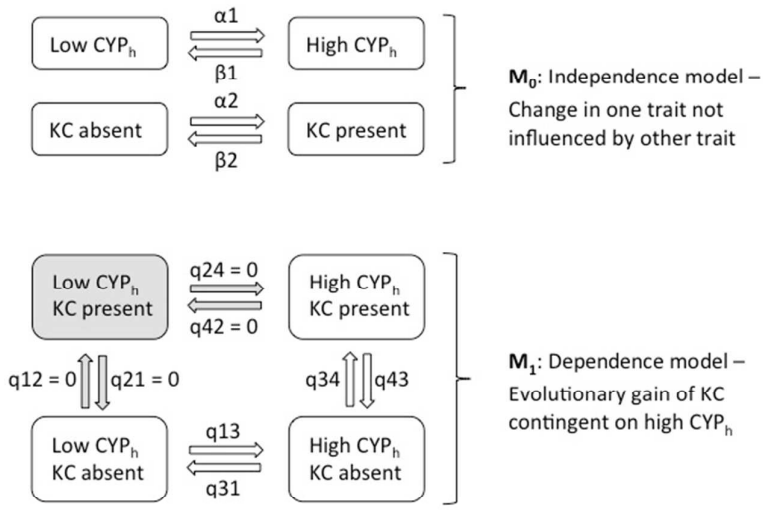


Figure 1. qRT-PCR quantification of *CYP2J19* expression in *Q. quelea* normalised against β-actin (N = 3) and shown on a logarithmic scale. Error bars represent SEM.



535 Figure 2. Alternative models for the evolution of hepatic *CYP2J19* expression (CYP_h) and red
536 ketocarotenoid pigmentation (KC) in weaverbirds. Arrows show evolutionary transition rates
537 that were either estimated (white) or restricted to zero (grey).

538

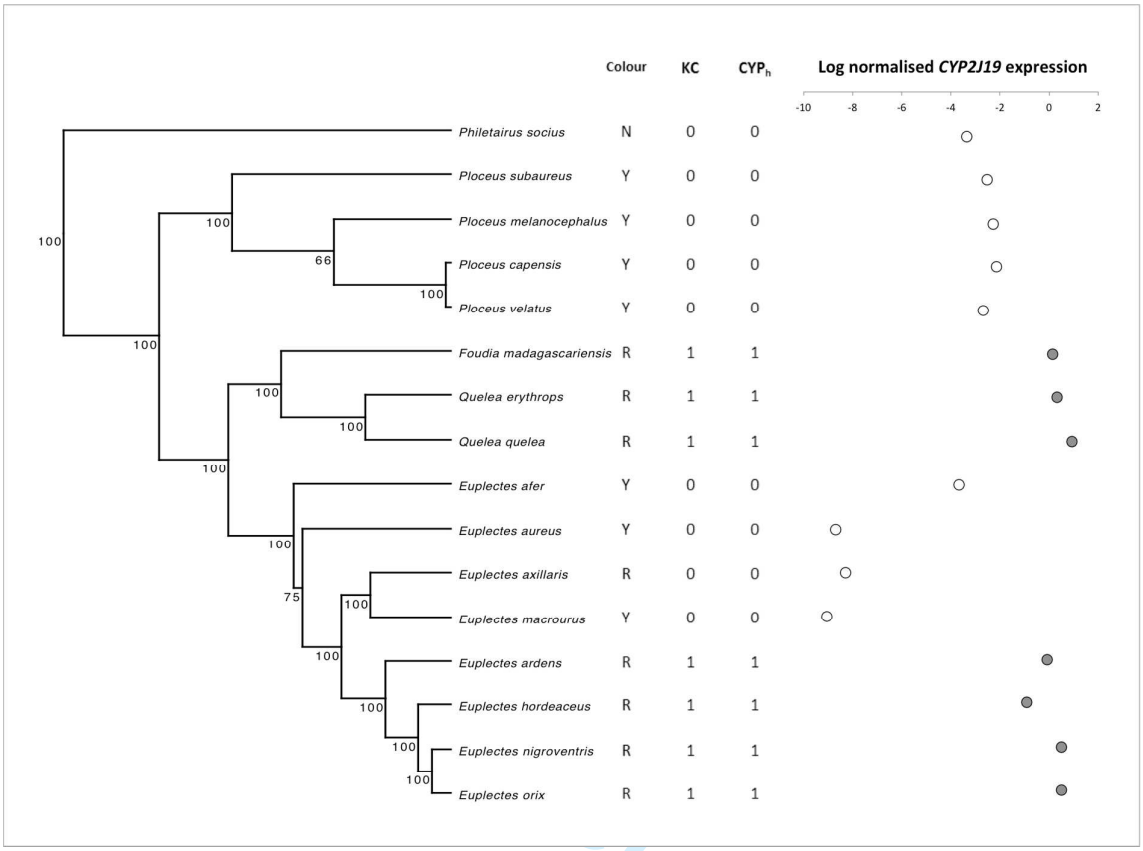


Figure 3: Hepatic expression of *CYP2J19* of 16 weaverbirds, in relation to phylogeny, coloration and ketocarotenoid presence. Gene expression was normalised against *β-Actin, GAPDH* and *HPRT1*, and $\log_{10}$-transformed. Expression levels of three species (*E. aureus, E. axillaris* and *E. macroura*) were undetectable after 50 PCR cycles. The phylogeny is a 50% majority-rule consensus (MRC) tree constructed in Mesquite 3.03 based on 10,000 trees downloaded from birdtree.org, numbers showing clade credibility (Bayesian posterior probability) in percent. Discrete scores of hepatic *CYP2J19* expression level (CYP_h: 0 = 'low', 1 = 'high', comprising white and grey dots on the continuous scale, respectively) and red ketocarotenoid pigmentation (KC: 0 = 'absent', 1 = 'present') used in evolutionary association tests are shown. The carotenoid-based coloration of the species (Red: 'R', Yellow: 'Y', Carotenoid absent: 'N'), is also shown.