



HAL
open science

Analysis of the genetic loci of pigment pattern evolution in vertebrates

Joel Elkin, Arnaud Martin, Virginie Courtier-Orgogozo, M. Emília Santos

► **To cite this version:**

Joel Elkin, Arnaud Martin, Virginie Courtier-Orgogozo, M. Emília Santos. Analysis of the genetic loci of pigment pattern evolution in vertebrates. *Biological Reviews*, 2023, 98 (4), pp.1250-1277. <10.1111/brv.12952>. <hal-04238685>

HAL Id: hal-04238685

<https://hal.science/hal-04238685v1>

Submitted on 2 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Analysis of the genetic loci of pigment pattern evolution in vertebrates

Joel Elkin^{1,*} , Arnaud Martin² , Virginie Courtier-Orgogozo³  and M. Emília Santos^{1,*} 

¹*Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK*

²*Department of Biological Sciences, The George Washington University, 800 22nd St. NW, Suite 6000, Washington, DC, 20052, USA*

³*Institut Jacques Monod, Université de Paris, CNRS, 15 Rue Hélène Brion, Paris 75013, France*

ABSTRACT

Vertebrate pigmentation patterns are amongst the best characterised model systems for studying the genetic basis of adaptive evolution. The wealth of available data on the genetic basis for pigmentation evolution allows for analysis of trends and quantitative testing of evolutionary hypotheses. We employed *Gephebase*, a database of genetic variants associated with natural and domesticated trait variation, to examine trends in how *cis*-regulatory and coding mutations contribute to vertebrate pigmentation phenotypes, as well as factors that favour one mutation type over the other. We found that studies with lower ascertainment bias identified higher proportions of *cis*-regulatory mutations, and that *cis*-regulatory mutations were more common amongst animals harbouring a higher number of pigment cell classes. We classified pigmentation traits firstly according to their physiological basis and secondly according to whether they affect colour or pattern, and identified that carotenoid-based pigmentation and variation in pattern boundaries are preferentially associated with *cis*-regulatory change. We also classified genes according to their developmental, cellular, and molecular functions. We found a greater proportion of *cis*-regulatory mutations in genes implicated in upstream developmental processes compared to those involved in downstream cellular functions, and that ligands were associated with a higher proportion of *cis*-regulatory mutations than their respective receptors. Based on these trends, we discuss future directions for research in vertebrate pigmentation evolution.

Key words: vertebrate pigmentation, development, evolutionary genetics, phenotypic variation, genetic variation, *cis*-regulatory mutations, coding mutations.

CONTENTS

I. Introduction	1251
(1) Vertebrate pigmentation	1252
(2) The role of <i>cis</i> -regulatory and coding mutations in evolution	1253
II. Methods	1253
(1) Literature curation in <i>Gephebase</i>	1253
(2) The vertebrate pigmentation data set	1254
(3) Assignment of new parameters	1254
(a) Clade	1254
(b) Pigment type	1254
(c) Phenotype category	1254
(d) Cellular function category	1256
(e) Molecular function category	1256
(4) Genetic hotspots	1257
(5) Statistical analysis	1257
(6) Correction for overrepresented genes	1257

* Authors for correspondence: J. Elkin (Tel.: +01223 336600; E-mail: joel.elkin416@gmail.com) and M.E. Santos (Tel.: +01223 336600; E-mail: es754@cam.ac.uk).

III. Results	1258
(1) Information on genetic variants associated with pigmentation is asymmetrically distributed across vertebrate clades	1258
(2) Variation in vertebrate pigmentation is associated most often with coding mutations	1258
(3) Study methodologies with less ascertainment bias exhibit a slightly higher proportion of <i>cis</i> -regulatory mutations	1258
(4) Changes in reporting of <i>cis</i> -regulatory mutations over time do not fully explain the disparity in frequency of reported mutation types	1259
(5) The dominance of coding mutations in the data set may be partly explained by studies on three genes involved in melanic pigmentation	1260
(6) The proportion of coding mutations is higher for studies of domesticated and intraspecific variation than for studies of interspecific variation	1260
(7) The proportion of coding mutations is higher in mammal and bird studies than in studies of teleosts, amphibians, and squamates	1261
(8) Types of phenotypic variation associated with coding and <i>cis</i> -regulatory mechanisms	1261
(9) Upstream developmental processes are associated with a higher proportion of <i>cis</i> -regulatory mutations	1263
(10) Transcription factors are associated with a higher proportion of <i>cis</i> -regulatory mutations than other molecular functions	1264
(11) Divergent pigmentation systems have fewer shared genetic hotspots	1265
IV. Discussion	1265
(1) Strong selection pressure might explain the higher proportion of coding mutations associated with vertebrate pigmentation diversity	1265
(2) Evolutionary timescales and taxonomic distance	1266
(3) Potential biases introduced by study systems	1266
(4) The importance of understanding the cellular and developmental basis of variation	1267
(5) What are the main knowledge gaps?	1267
V. Conclusions	1268
VI. Acknowledgements	1269
VII. References	1269
VIII. Supporting information	1276

I. INTRODUCTION

One of the central goals of evolutionary biology is to understand the genetic basis of organismal diversity. Determining which genes and mutations underlie adaptive traits is essential for understanding how evolutionary forces shape organismal variation (Barrett & Hoekstra, 2011). In this regard, vertebrate pigmentation is a powerful system for integrating historically disparate fields of evolutionary research, and ultimately identifying the genetic mechanisms underlying evolutionary change (Cuthill *et al.*, 2017; Hubbard *et al.*, 2010; Kronforst *et al.*, 2012; Orteu & Jiggins, 2020).

Vertebrate pigmentation offers a diverse array of phenotypes, with intra- and interspecific variation ranging from whole-body colouration to highly localised pattern alterations (Fig. 1). The adaptive significance of pigment patterns can be attributed to multiple distinct selection pressures, including thermoregulation, camouflage, aposematism, sexual display, and ultraviolet protection (Protas & Patel, 2008). The rapid evolution of vertebrate pigmentation patterns, and their evolutionary significance, allows for both mapping of individual mutations to their resultant phenotypes, and inference of the evolutionary pressures driving the selection of those phenotypes. There have been many studies identifying loci associated with a wide range of pigmentation phenotypes in different vertebrate taxa. Analysis

of these individual case studies may reveal broad trends underlying pigment pattern evolution and help to inform the future direction of vertebrate pigmentation research.

Here, we first outline the basic biology of vertebrate pigmentation, as well as the key differences between different vertebrate clades. We then utilise a data set of vertebrate pigmentation literature to analyse trends in the underlying genetics. For this purpose, we used *Gephebase* (<https://www.gephebase.org/>) – a knowledge base dedicated to the compilation of literature on genes and mutations underlying natural and domesticated organismal variation in eukaryotes (Courtier-Orgogozo *et al.*, 2020). *Gephebase* gathers published data on evolutionarily relevant mutations, with each entry representing a causal association between an allelic change at a given locus and trait variation between individuals or species. Each entry includes information relating to the species/population, the type of trait, the gene, the nature of the mutation(s), whether they represent null alleles, and the study that identified it. Using the vertebrate pigmentation entries in this database, we focused on the relative abundance of *cis*-regulatory and coding mutations contributing to vertebrate pigmentation evolution. Finally, we discuss the direction of future vertebrate pigmentation research and the role to be played by recent model systems and innovative approaches.



Fig. 1. Examples of vertebrate colour pattern diversity. (A) Sympatric colour morphs of the European common wall lizard (*Podarcis muralis*) (Andrade *et al.*, 2019). The colour morphs differ in a range of key morphological, physiological, and behavioural traits. Photographs courtesy of Pedro Andrade and Miguel Carneiro. (B) Allopatric morphs of harlequin poison frogs (*Oophaga histrionica* complex). Amphibian colour patterns are extremely under-represented in pigmentation evolutionary genetic studies. Photographs courtesy of Roberto Marquez. (C) Two species of Lake Victoria cichlids. *Haplochromis sawagei* (top) and *Pundamilia nyererei* (bottom) showing differences in horizontal stripes and vertical bars (Kratochwil *et al.*, 2018). Photographs courtesy of Claudius Kratochwil. (D) Different species of *Peromyscus* mice showing differences in colour phenotypes. *P. maniculatus nubiterrae* (top left, image credit Evan P. Kingsley), *P. polionotus phasma* (top right, image credit J.B. Miller), *P. polionotus sumneri* (bottom left, image credit Nicole Bedford) and *P. gossypinus* (bottom right, image credit Nicole Bedford). Differences in melanic coat colouration correlate with the colour of the background and evolved as a result of strong predation. Photographs adapted from Bedford & Hoekstra (2015). (E) Red and black head-colour polymorphism in the Gouldian finch (*Erythrura gouldiae*), morphs display differences in aggression and reproductive success (Toomey *et al.*, 2018). Photograph courtesy of Ricardo Jorge Lopes and Miguel Carneiro.

(1) Vertebrate pigmentation

In vertebrates, most colour patterns derive from specialised pigment cells which produce either pigments or reflective structures. These cells arise from the migratory neural crest

cell population, which emerges from the dorsal neural tube during early vertebrate development (Lapedriza, Petratou & Kelsh, 2014). Neural crest cells then delaminate and undergo some of the longest migrations of any embryonic cell type to

give rise to multiple derivatives such as neurons and glia of the peripheral nervous system, smooth muscle, craniofacial cartilage and bone, and pigment cells (Simões-Costa & Bronner, 2015).

Across vertebrate clades there is considerable diversity in the cell types producing colours. In fishes, amphibians, and non-avian reptiles there are multiple distinct classes of pigmented and structurally coloured cells, called chromatophores. These classes contain different combinations of pigments and/or reflective structures, and therefore exhibit different ranges of colours. Across these clades up to at least nine chromatophore classes are recognised. The most common chromatophore classes that utilise pigments are melanophores (brown/black melanin pigments) and xanthophores (yellow/orange carotenoid and pteridine pigments). These cells produce pigment-based colour *via* the deposition of their respective pigment molecules, which selectively absorb specific wavelengths of light. By contrast, structural colouration results from the presence of light-interfering structures, such that structural colour is variable depending on the angle from which it is viewed. Iridophores are structural chromatophores, which appear silvery or blue due to the arrangement of layered purine platelets of variable size, shape, and arrangement (Parichy, 2021; Singh & Nüsslein-Volhard, 2015).

Rarer chromatophore classes include red erythrophores and blue cyanophores, as well as two distinct classes of white leucophores with different regulatory profiles, developmental origins, and chemical compositions (Goda & Fujii, 1995; Huang *et al.*, 2021; Lewis *et al.*, 2019). Each chromatophore class present in an organism goes through extensive cell movements and cell–cell interactions to form a final colour pattern. Variation in the abundance, combinations, and arrangements of chromatophores generates the intricate and diverse colour patterns seen in fishes, amphibians, and reptiles.

By contrast, mammals and birds have independently lost most pigment cell diversity and mainly retain only one pigmentary cell type, the melanocyte (equivalent to the melanophore) (Kelsh *et al.*, 2009). Melanocyte differentiation and development remains highly conserved among vertebrates, but differences exist. Unlike melanophores, melanocytes produce two melanin pigment types in different shades – brownish-black eumelanin, and reddish pheomelanin. The ability to switch melanogenesis between eumelanin and pheomelanin production is specific to birds and mammals (McNamara *et al.*, 2021). Mammals and birds develop complex pigmentation patterns by temporally and spatially regulating melanogenesis switching, as opposed to using different chromatophore classes for different colours. Furthermore, while other vertebrates retain pigments within the respective chromatophore, in birds and mammals melanin is secreted from melanocytes into the skin or the integumentary appendages, such as feathers and hairs (McNamara *et al.*, 2021). Birds additionally exhibit an array of yellow, orange and red colours due to the processing and accumulation of dietary carotenoid pigments (Toews, Hofmeister & Taylor, 2017). Finally, birds also display

structural colouration arising from reflective nano-structures in both the skin and plumage (Saranathan & Finet, 2021).

(2) The role of *cis*-regulatory and coding mutations in evolution

The extent to which evolution is predictable on the level of genetic variation is a long-standing topic in evolutionary biology. One important question concerns the relative prevalence of *cis*-regulatory and coding sequence mutations (Carroll, 2008; Hoekstra & Coyne, 2007; Martin & Courtier-Orgogozo, 2017; Martin & Orgogozo, 2013; Stern & Orgogozo, 2009, 2008). It has been hypothesised that *cis*-regulatory mutations are more likely to contribute to evolutionary changes in morphology (Carroll, 2008). One reason for this is that *cis*-regulatory mutations are expected to have fewer pleiotropic effects when compared to coding mutations, mostly due to their highly modular nature conferring tissue specificity, thus affecting gene expression patterns without changing protein function. Conversely, protein coding sequences are less modular, and non-synonymous mutations are expected to affect protein function across every cell and tissue in which it is expressed. However, for the past three decades, case studies have shown that both types of mutations can contribute to evolutionary change. Thus, trying to argue for the existence of a dichotomy in the relative frequency of these types of genetic change is too simplistic. Instead, we should strive to identify the contexts in which one type of mutation is favoured over the other (Stern & Orgogozo, 2008). For example, are different mutations associated with particular aspects of trait variation (e.g. colour *versus* pattern)? Do the cellular and developmental processes underlying a trait influence the nature of the mutations that can affect it? Do we see a shift in the role played by different types of mutation over evolutionary time?

Herein, we tackle these questions by examining the relative distribution of *cis*-regulatory and coding mutations across vertebrate pigmentation evolution, and how the relative frequencies of these mutation types associate with other factors, such as loci mapping strategy, type of trait variation or types of genes involved. Moreover, by analysing the types of mutation associated with the evolution of vertebrate pigmentation, we identify trends in the field and discuss possible directions for future research.

II. METHODS

(1) Literature curation in *Gephebase*

Gephebase compiles pairs of alleles in association with pairs of phenotypic states (a genetic variation causing a phenotypic variation is called a ‘gophe’ for brevity). A full description of the database is provided in Courtier-Orgogozo *et al.* (2020). Data currently cover all eukaryotes with relevant publications, with a focus on traits of evolutionary rather than

clinical relevance. This includes variations that have been artificially selected by breeders ('Domesticated' data set), or subject to experimental evolution under laboratory-controlled selective pressures. Data are manually curated by a small team of researchers, using key words as well as Pubmed/Google Scholar citations to identify newly published studies. The current triage system gives priority to genes identified by forward genetics – quantitative trait locus (QTL) mapping, and genome-wide association studies (GWAS) with single-gene resolution and reasonable supporting evidence for causality (if not for the variant, at least for the gene). Genes identified by candidate gene approaches without mapping are also included when there is additional functional evidence for the causality of the mutation. *Gephebase* is up to date and gathers all relevant papers published until 2017. Since 2017, data curation efforts have mainly focused on colour variation in vertebrates (trait category = 'Coloration'). The download of *Gephebase* data on 15 August 2022 (see online supporting information, File S1, tab: All Vertebrates), which was used for the present study, represents a compilation of the genes and mutations contributing to vertebrate colouration up to August 2022.

(2) The vertebrate pigmentation data set

To examine trends in colour pattern evolution, we utilised *Gephebase* as a resource for exploring genotype–phenotype relationships. We formed a working data set by selecting every *Gephebase* entry pertaining to the trait category 'Coloration', and further filtered those entries by removing all entries pertaining to non-vertebrate species (accessed 15.08.2022). In total we retrieved 389 entries, with each entry representing a mutation at a single locus. Each mutation was present in only one gene for the organism in which it was identified, and there were a total of 68 unique genes identified across 103 vertebrate species. For each entry, we examined five parameters defined by *Gephebase*: Gene ID, Taxonomic status, Study methodology, Aberration type, and Molecular type (Courtier-Orgogozo *et al.*, 2020). We further defined five additional parameters detailed below: Clade, Pigment type, Phenotype category, Cellular function category, and Molecular function category (supporting information, File S1, tabs: All Vertebrates, GO Terms by Gene).

Table 1. Number of entries in the vertebrate pigmentation *Gephebase* data set in 2022 (see supporting information, File S1, tab: All Vertebrates) according to vertebrate clade and taxonomic status. One entry in the *Gephebase* database corresponds to genetic variation at a given gene that has been found to contribute to pigmentation evolution in a given taxon.

Gephebase entries	Domesticated	Intraspecific variation	Interspecific variation	Total by clade
Teleosts	6	11	12	29
Amphibians	1	1	0	2
Squamates	1	9	0	10
Birds	57	19	6	82
Mammals	182	82	2	266
Total by taxonomic status	247	122	20	389

(3) Assignment of new parameters

(a) Clade

Each vertebrate represented in the data set belonged unambiguously to one of five clades with distinct pigmentation biology: amphibians, teleosts, squamates, birds and mammals (Table 1). We divided entries into these clades to compare trends in their different pigmentation systems.

(b) Pigment type

In total, five pigment types were represented in the data set: biliverdin, carotenoids, pteridines, melanins, and psittacofulvins. Additionally, four entries referred to structural colouration. This is a short list of pigments, but few studies identified the specific pigment molecules associated with a phenotype. For example, a phenotype may clearly implicate a change in the synthesis or distribution of melanin pigments, without distinguishing between pheomelanin and eumelanin. Hence, these two types of pigment are grouped in the melanin category. Finally, there were 13 entries in which a single genetic variant was associated with changes to multiple pigments. These entries were assigned to each relevant pigment type category.

(c) Phenotype category

We empirically assigned to each entry one of seven phenotype categories (Table 2) in order to group similar organismal colour loss/tuning phenotypes, and analysed trends in how they arise. Each entry was considered independently, and assigned the category that best described the organismal phenotype based on the original study. We accounted for both the visible phenotype in terms of pigment distribution and quantity, as well as the molecular mechanisms underlying the phenotype. For example, amelanism was distinguished from white spotting on the basis of whether the phenotype resulted from a loss of melanin synthesis, or impairment of melanocyte migration/differentiation.

A total of 54 entries could not be assigned a category unambiguously. These included phenotypes involving global developmental changes that indirectly affect pigment pattern, for instance, a *cyp19a1* mutation that results in female-like plumage patterns in male chickens due to ectopic

Table 2. Summary of phenotype categories: assignment criteria for each phenotype category, and the total number of entries in the data set for which that category was assigned. SNP, single nucleotide polymorphism.

Category name	Criteria for assignment	Number of entries	Example
Colour loss <i>versus</i> colour tuning			
Pheomelanism	A phenotype exhibiting increased abundance of pheomelanin relative to eumelanin, through changes to melanogenic switching or production of either melanin type. This includes phenotypes in which eumelanin synthesis is absent or reduced such that pheomelanin is relatively more prevalent.	87	A missense mutation in <i>mc1r</i> is associated with the red phenotype in domesticated donkeys (<i>Equus asinus</i> ; Abitbol <i>et al.</i> , 2014).
Eumelanism	A phenotype exhibiting increased abundance of eumelanin relative to pheomelanin, through changes to melanogenic switching or production of either melanin type. This includes phenotypes in which pheomelanin synthesis is absent or reduced such that eumelanin is relatively more prevalent.	83	A loss-of-function SNP in <i>asip</i> is associated with ‘black panther’ phenotype melanistic leopards (<i>Panthera pardus</i> ; Schneider <i>et al.</i> , 2012).
Dilution	A phenotype in which both eumelanin and pheomelanin synthesis is reduced (but not lost) across one or several tissues.	22	A melanophilin splice-site SNP leads to exon skipping and coat colour dilution in domesticated rabbits (<i>Oryctolagus cuniculus</i> ; Lehner <i>et al.</i> , 2013).
Amelanism	A phenotype characterised by a loss of melanin production across one or several tissues.	62	Three separate <i>tyrosinase</i> coding mutations are associated with oculocutaneous albinism in three species of frog (<i>Pelophylax nigromaculatus</i> , <i>Glandirana rugosa</i> & <i>Fejervarya kavamurai</i> ; Miura <i>et al.</i> , 2018).
Loss of carotenoid colouration	A phenotype resulting from the inability to take up or process dietary carotenoids causing total loss of carotenoid-based pigmentation.	3	A splice-site mutation in <i>scarb1</i> inhibits cellular carotenoid uptake, leading to a total loss of carotenoid pigmentation in white canaries (<i>Serinus canaria</i> ; Toomey <i>et al.</i> , 2017).
Change in carotenoid colouration	A phenotype resulting from changes to the relative or overall abundance of carotenoid-based pigments in the body.	11	Gain of expression of a carotenoid ketolase, <i>cyp2j19</i> , is responsible for processing yellow carotenoids to produce red ketocarotenoids in red canaries (<i>Serinus canaria</i> ; Lopes <i>et al.</i> , 2016).
White spotting	A phenotype resulting from impairment of migration and/or differentiation of melanocytes or their precursors, causing patchy regions of white tissue. Differentiated from amelanism through molecular mechanism rather than visual appearance.	67	White spotting phenotypes in domestic horses are associated with a large number of distinct mutations identified in <i>kit</i> (<i>Equus caballus</i> ; Haase <i>et al.</i> , 2015a).
Colour shift <i>versus</i> pattern alteration			
Colour shift	Any phenotype in which organismal colour shifts either locally or globally, without impacting the spatial boundaries of its existing patterning. Includes phenotypes in which an organism’s existing pattern is overridden, e.g. whole-body eumelanism.	222	In Japanese quail (<i>Coturnix japonica</i>), a frameshift deletion in <i>asip</i> is associated with the recessive black phenotype, in which feather colour is darkened across the whole body without altering pattern boundaries (Hiragaki <i>et al.</i> , 2008).
Pattern alteration	Any phenotype in which the spatial boundaries of an organism’s colour pattern change, including the formation of new boundaries e.g. establishment of a pattern in a previously uniform region. Does not include phenotypes in which an existing pattern changes colour without change to the pattern boundary.	129	The convergent evolution of horizontal melanic stripes in different cichlid species is associated with regulatory changes in <i>agouti-related peptide 2</i> , and expression levels are predictive of stripe presence across multiple cichlid radiations (Kratochwil <i>et al.</i> , 2018).

oestrogen synthesis in the skin (Li *et al.*, 2019b). Likewise phenotypes involving pigment changes without a clear causative mechanism were not categorised. For instance, a blue egg-shell phenotype in chickens is associated with a *slco1b3* allele that is overexpressed in the shell gland, but the mechanism by which *slco1b3* affects shell pigmentation remains unclear

(Wang *et al.*, 2013). Removing ambiguous cases left a total of 335 assignments, of which 306 were *cis*-regulatory or coding. Additionally, categories were assigned relative to the ancestral state, so that, for instance, eumelanism refers to a mutation leading to an increase in eumelanin pigmentation.

We additionally assigned each category to either colour shift or pattern alteration (Table 2), on the basis of whether the phenotype affected colour or spatial pattern boundaries. Each entry was considered independently, and these assignments were considered separately from the previous phenotype categories. Phenotypes involving a loss of patterning (for example whole-body albinism resulting in no visible pattern boundaries) were considered colour shifts.

A total of 38 entries could not be assigned a category unambiguously. These included phenotypes for which the distinction between colour and pattern changes was unclear, particularly in the case of domesticated breeds where the ‘ancestral’ pigmentation pattern is poorly defined. For instance, an *asip* mutation associated with a uniform coat colour in dogs is found in breeds that otherwise exhibit various coat phenotypes both patterned and uniform, making comparison to an ancestral state unclear. Removing ambiguous cases left a total of 351 assignments, of which 317 were *cis*-regulatory or coding.

(d) Cellular function category

In order to analyse genes associated with different developmental and cellular roles, we assigned each gene in the data set a functional category, reflecting its role in the development of the phenotype (Table 3). Unlike phenotype categories, these

assignments were based on the gene rather than on individual entries. Entries were assigned empirically, based on a literature review of the genes identified by each study, and the phenotypic function(s) they have been implicated in. As with phenotype categories, each gene was assigned with high certainty and little ambiguity. A total of four genes, comprising seven entries, were removed from this analysis, as their cellular functions specifically with regard to pigmentation biology were unclear. These were *cyp19a1*, *spint1*, *ejf2s2*, and *goldentouch*, an undescribed gene. These genes had one entry each.

(e) Molecular function category

We assigned protein molecular function categories non-empirically by using the gene ontology (GO) terms associated with each gene in the data set. We employed the EBI QuickGO Mouse Slimmer in order to identify parent GO terms associated with the sets of child terms associated with each gene. This slimmer was selected as being the most representative of the distribution of GO terms within vertebrate clades. We then narrowed the set of GO terms to include only those associated with the ‘Molecular function’ GO tree (see supporting information, File S1, tab: GO Assignments). We additionally manually combined a number of closely related GO terms in order to reduce the number of potential

Table 3. Summary of cellular functional categories: assignment criteria, and the total number of entries in the data set to which that category was assigned. Each gene was assigned to only one functional category.

Cellular functional category	Assignment criteria	Number of unique genes	Number of entries	Example
Pigment cell fate specification	Contributes to determining fate specification of pigment cell precursor lineages	10	21	Transcription factor <i>sox10</i> is required for specifying melanocyte fate in neural crest cells (Harris <i>et al.</i> , 2010)
Pigment cell precursor survival/migration	Necessary for the survival and migration of pigment cells post-specification	15	82	Post-specification survival and migration of melanoblasts is dependent on <i>kit</i> and <i>kitlg</i> regulation (Mort, Jackson & Patton, 2015)
Cellular interactions	Involved in signalling or other cellular interactions, either between or within pigment cell classes	4	7	In teleosts, <i>csf1a</i> encodes a cytokine that facilitates signalling between chromatophore classes, enabling pigment cell interactions that govern pattern formation (Patterson & Parichy, 2013)
Melanosome formation	Contributes to the formation or maintenance of melanosomes or specialised melanin vesicles	8	48	The transporter protein encoded by <i>slc45a2</i> contributes to regulating melanosomal pH during maturation of melanosomes (Le <i>et al.</i> , 2020)
Pigment biosynthesis	Role in directly synthesising or processing pigments or their direct precursors; includes regulating melanogenic switching	23	211	Tyrosinase is an essential enzyme in melanin biosynthesis; required for multiple reactions in the synthesis pathways of both eumelanin and pheomelanin (del Marmol & Beermann, 1996)
Pigment trafficking/localisation	Role in inter- or intracellular localisation of pigments, whether as a direct transporter or as a trafficking protein	4	16	Intracellular melanosome transport is dependent on a molecular motor complex incorporating melanophilin; <i>mlph</i> mutants exhibit melanosome clustering around the nucleus (Matesic <i>et al.</i> , 2001)

categories (see Table 4; supporting information, File S1, tab: Molecular Function Summary). In total there were 89 unique assignments of parent GO categories across the data set. All but two of the 68 unique genes in the data set generated at least one assignment. The only two exceptions were *goldentouch*, a previously undescribed gene identified in Midas cichlids with no known molecular function, and *mlana*, a gene implicated in melanosome biogenesis. Only one molecular function assignment (GO:0005515 protein binding) is assigned to *mlana*, which is not included in the QuickGO Mouse Slimmer. After combining closely related terms, there were 10 parent GO terms assigned. Then, we removed all categories with fewer than five unique gene assignments, of which there were three: RNA binding, enzyme regulator activity, and lipid binding. Thus we ended up with seven molecular function categories for comparison. The DNA binding GO category was the result of combining two GO terms: DNA binding and transcription factor activity. Although there are distinct biological differences between a gene displaying DNA binding activity and one acting as a transcription factor, in the *Gephebase* data set there was nearly complete overlap between these terms. Only one gene was tagged as DNA binding without also being tagged for transcription factor activity: epidermal growth factor receptor *egfr*. All other genes were tagged either with both GO terms, or with neither.

Table 4. The 10 molecular function (parent gene ontology, GO) categories assigned to data set entries, the total number of genes to which each GO category was assigned and corresponding total number of entries, the number of *cis*-regulatory entries to which each was assigned and corresponding *cis*-regulatory proportion, and the child GO terms that were combined into the category where relevant. For analysis we removed categories with fewer than five unique assignments (RNA binding, enzyme regulator activity, and lipid binding; see Section II.3).

Molecular function category	Unique genes assigned	Total entries assigned	<i>Cis</i> -regulatory entries	<i>Cis</i> -regulatory proportion	Child GO terms	GO term ID
Catalytic activity	17	121	15	0.146	Transferase activity, oxidoreductase activity, hydrolase activity, ligase activity	GO:0003824
Signalling receptor binding	14	66	25	0.463	N/A	GO:0005102
Transporter activity	13	44	5	0.122		GO:0005215
DNA binding	12	31	17	0.630	DNA binding, transcription factor activity	GO:0003677
Signalling receptor activity	10	149	9	0.067		GO:0038023
Protein binding	8	39	9	0.257	Cytoskeletal protein binding	GO:0005515
Carbohydrate derivative binding	6	54	8	0.186		GO:0097367
Enzyme regulator activity	4	4	1	0.5		GO:0030234
RNA binding	3	4	2	1		GO:0003723
Lipid binding	2	2	0	0		GO:0008289
Totals	89	514	91	0.177		

(4) Genetic hotspots

We investigated the overlap of genes between pigmentation systems by examining the number of entries corresponding to shared genes. For each clade (excluding amphibians due to lack of data), we listed all the genes identified in that clade, and then calculated the total number of entries corresponding to those same genes in each of the other three clades. We normalised each of these figures by dividing by the total number of entries in that clade, in order to account for the disparity in entries between clades.

(5) Statistical analysis

For each analysis the *cis*-regulatory proportion was fitted as a binomial logistic regression model and *post hoc* pairwise comparisons were performed with Tukey correction using the `glht` function in the `multcomp` package for R. Direct comparisons between categories or groups of categories were performed using Fisher's exact test, again at $P < 0.05$.

(6) Correction for overrepresented genes

For each analysis we compared the total numbers of *cis*-regulatory and coding mutations in each category. However, some particularly well-characterised genes contribute a

disproportionate number of entries to the data set. To assess whether this affected our results, we repeated all analyses by treating each unique gene as a single data point, with a proportion of *cis*-regulatory or coding mutations calculated for that gene's entries in a particular category. For instance, for *asip* in birds, there were three entries of *cis*-regulatory mutations out of five entries in total, so *asip* was treated as a single data point with a value of 0.6 for *cis*-regulatory mutations in birds; in mammals, *cis*-regulatory mutations in *asip* instead had a value of 0.32. Each plot shows the proportions of *cis*-regulatory and coding mutations by category, manually scaled to sum to 1 to allow comparison with other figures. *Post hoc* statistical tests were not performed on these plots.

III. RESULTS

(1) Information on genetic variants associated with pigmentation is asymmetrically distributed across vertebrate clades

To examine the representation of different pigmentation systems across vertebrate evolutionary studies, we looked at the number of entries in *Gephebase* across five clades: teleost fishes, amphibians, squamates, birds and mammals. All 389 entries belonged to one of these clades. The significant majority (68.3%, $N = 266$) of entries were for mammals, followed by birds (21.1%, $N = 82$), teleosts (7.5%, $N = 29$), squamates (2.6%, $N = 10$) and amphibians (0.5%, $N = 2$). Clades in which multiple distinct chromatophore classes are found thus represent only $\sim 10\%$ of our data set, with the majority of data from clades with only melanocytes and no other pigment cell classes. This taxonomic bias is attributable in part to how this field originated and evolved. The disproportionate representation of mammalian studies is likely a reflection of the extensive characterisation of mouse coat colour genetics, which has been studied for more than a century (Hoekstra, 2006). As a result, many of the key components of the melanin biosynthesis pathway were first identified in mice, and these components are often selected for investigation in other mammalian and bird systems.

(2) Variation in vertebrate pigmentation is associated most often with coding mutations

We examined the relative prevalence of protein coding and *cis*-regulatory mutations across the data set. *Cis*-regulatory mutations, and coding sequence mutations together account for 348 entries (89.5%) in the full data set, with coding mutations representing the majority of entries (272, 69.9%) (Fig. 2A). The remaining classes of genetic variation were gene amplification, gene loss, intronic mutations, and unknown (when the gene was identified but the exact mutation(s) involved were not specified) (Fig. 2A). This dominance of coding mutations is in contrast to the prediction that *cis*-regulatory mutations are more likely to generate phenotypic change (Carroll, 2008). However, the observed

pattern may reflect the relative ease of identifying coding mutations compared with *cis*-regulatory changes (Stern & Orgogozo, 2008).

(3) Study methodologies with less ascertainment bias exhibit a slightly higher proportion of *cis*-regulatory mutations

Three methodology categories are represented in the data set: candidate gene, linkage mapping, and association mapping (Courtier-Orgogozo *et al.*, 2020). Of these, candidate gene studies are likely to involve the highest ascertainment bias and may be expected to identify a higher proportion of coding mutations. Many candidate gene studies historically focused on identifying amino acid changes within coding regions in new species based on previous findings in other species. However, they can facilitate comparisons across broad taxonomic distances compared with other mapping approaches. By contrast, both linkage and association mapping studies start with a phenotypic difference and attempt to link it to a sequence change, or at least to a locus interval. Linkage mapping involves crossing populations to generate recombinant hybrids, and thus can only map variation between closely related organisms that are interfertile. Such studies are less likely to suffer from ascertainment bias than candidate gene approaches in their detection of causative loci, depending on their resolution and mapping coverage. Finally, association mapping approaches involve identification of significant associations within large, heavily intermixed populations, typically with minimal ascertainment bias as a result. To account for potential differences between methodologies we plotted the proportions of *cis*-regulatory and protein coding mutations for each method (Fig. 2B).

As expected, candidate gene approaches identified the lowest proportion of *cis*-regulatory mutations (12.5%). For linkage mapping studies, this value was 25.3%, and for association mapping it was higher still (46.2%). Thus all three of these methodology categories result in more reported coding than *cis*-regulatory mutations (Fig. 2B), and regardless of the method used, protein coding mutations appear to be involved in the majority of investigated vertebrate pigmentation variation. An effect of study methodology may be seen in that the two methodologies with a lower ascertainment bias appear to report more causal *cis*-regulatory mutations.

As a cautionary note, all methodologies will tend to favour the identification of coding mutations. For example, once a candidate region is identified by linkage mapping or association mapping, it is easier to find coding mutations contributing to the trait of interest than *cis*-regulatory mutations: based on the genetic code it is simpler to identify mutations that disrupt the amino acid sequence of the encoded protein than it is to predict *cis*-regulatory effects based on sequence alone. Furthermore, validation of *cis*-regulatory mutations usually requires the use of reporter constructs and transgenic animals, and thus is more time-consuming than the validation of coding mutations which often only involves *in vitro* assays.

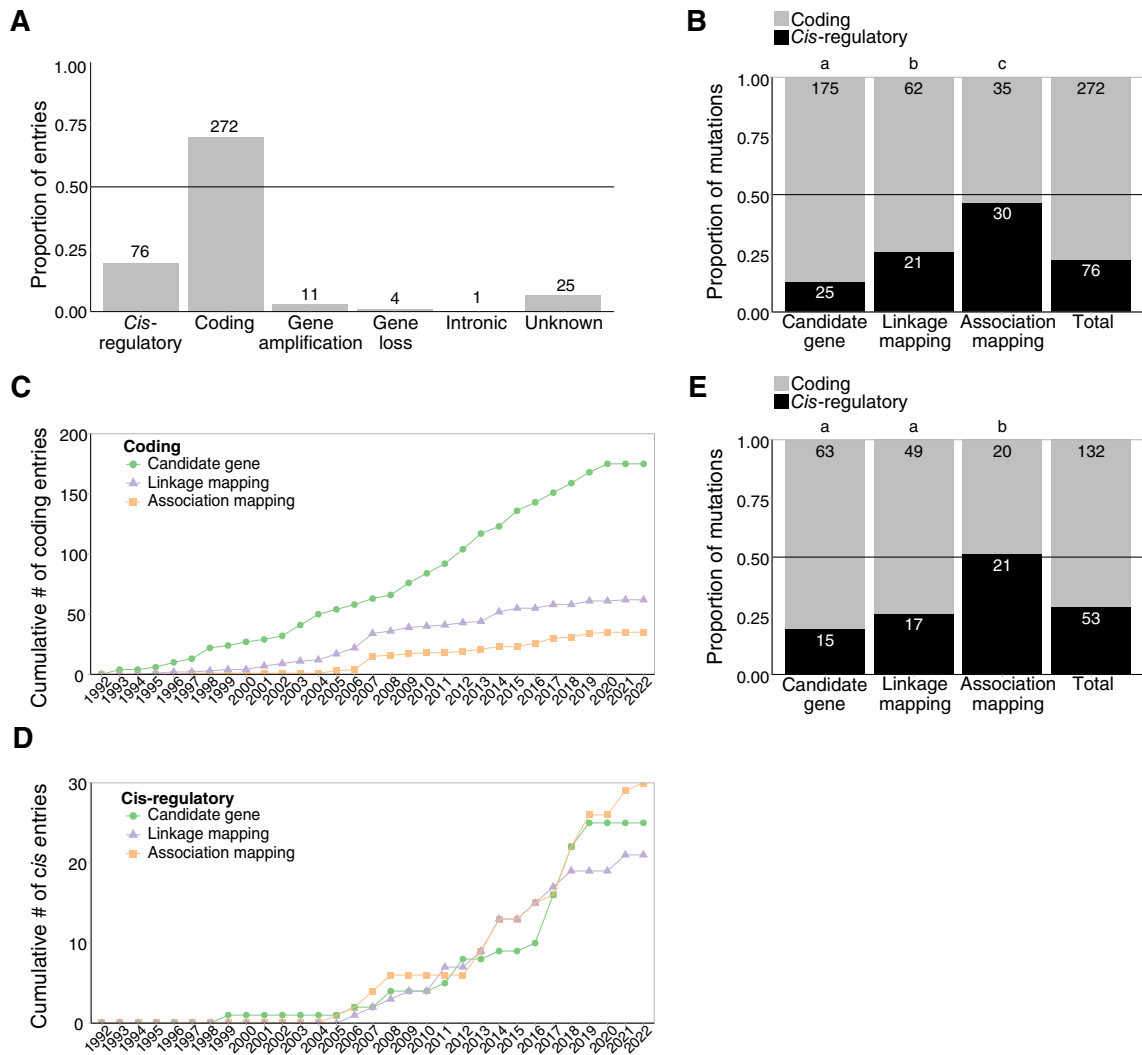


Fig. 2. (A) Relative proportions of all mutation types identified in our data set. (B) Relative proportions of *cis*-regulatory and coding mutations associated with each study methodology, and in the total data set. (C, D) Cumulative number of reported coding mutations (C) and *cis*-regulatory mutations (D) for each study methodology over the time period included in the *Gephebase* data set. (E) *Cis*-regulatory and coding mutations associated with each study methodology, with *mc1r*, *asip* and *kit* entries removed. In B and E, letters above columns indicate results of *post hoc* pairwise comparisons with Tukey correction ($P < 0.05$); categories that are not significantly different share a letter. In A, B and E, the numbers above/within the bars represent the number of entries in that category.

(4) Changes in reporting of *cis*-regulatory mutations over time do not fully explain the disparity in frequency of reported mutation types

The *Gephebase* data set includes vertebrate pigmentation studies published between 1993 and 2022 (Courtier-Orgogozo *et al.*, 2020). Due to the increasing availability of genomic resources facilitating linkage and association mapping in a broader range of model systems (Kratochwil & Meyer, 2015), we expected reporting of *cis*-regulatory mutations to increase over time. We therefore examined the cumulative number of *cis*-regulatory and coding mutations over the time period covered by the data set, for each study methodology. We used this approach as the low numbers of entries in

each individual year made non-cumulative comparisons misleading – for example when there were years with no entries for a particular study methodology. Entries for *cis*-regulatory mutations first appeared in the data set later than coding entries (Fig. 2C, D). Prior to 2005 only a single *cis*-regulatory mutation was identified, compared with 63 entries for coding mutations. However, since 2005, both *cis*-regulatory and coding mutations have been added at approximately constant rates for each study methodology, with coding mutations added at a higher rate for each methodology. Thus, the ratio of *cis*-regulatory to coding mutations remains roughly constant over time, and coding mutations constitute the majority of mutations identified in each year that this data set covers.

We also did not observe an appreciable change in the relative prevalence of linkage mapping or association mapping studies (Fig. 2C, D). Taken together, while there is an overall trend for increased reporting of *cis*-regulatory mutations with time, differences in the relative prevalence of different study methodologies are insufficient to explain the preponderance of coding mutations in our data set. It remains possible that there is a discovery bias towards the identification of coding mutations, or alternatively, that evolutionary variations in vertebrate pigmentation do indeed involve coding mutations more often.

(5) The dominance of coding mutations in the data set may be partly explained by studies on three genes involved in melanic pigmentation

Historically, most evolutionary studies have focused on mammalian and bird melanic pigmentation systems, presumably due to the large number of studies on the genetics of mouse coat colour. Therefore, we explored if the observed preponderance of protein coding mutations could result from the overrepresentation of melanic pigmentation case studies. Mammal and avian melanic pigment patterns are the result of dynamic switching between the synthesis of two types of melanin. The switch between black eumelanin and reddish-brown pheomelanin is controlled by the melanocortin 1 receptor (MC1R) encoded by *mc1r* (García-Borrón, Sánchez-Laorden & Jiménez-Cervantes, 2005). With 88 entries (22.6% of the data set) *mc1r* is the most common gene in our data set. The next most represented gene is *kit*, which encodes a receptor tyrosine kinase essential for the survival and migration of melanocyte/melanophore precursors (Hou, Panthier & Arnheiter, 2000) and represents 12.6% (49 entries) of the data set. The third most represented gene is *asip*, which codes for agouti signalling protein (ASIP). This protein acts as an antagonist to MC1R; in the absence of ASIP, MC1R initiates a signalling cascade that results in eumelanin production, while when ASIP is present the melanocyte reverts to pheomelanin synthesis (García-Borrón *et al.*, 2005). There are 43 *asip* entries in our data set (11.1%).

The overrepresentation of these three genes in *Gephebase* reflects their relevance in mammalian and other vertebrate pigmentation systems, and particularly their frequent selection as candidate genes for further study. To remove any potential bias resulting from a focus on these genes, we re-analysed the data with *mc1r*, *kit*, and *asip* removed. With these entries removed from the data set, the overall proportion of *cis*-regulatory mutations increased from 21.8% to 28.6%, due to the prevalence of coding mutations in *mc1r* entries in particular (Fig. 2E). Similarly, the proportions of *cis*-regulatory mutations in all three methodology categories increased (Fig. 2E). Coding mutations remain the most reported for candidate gene and linkage mapping methodologies, but for the association mapping category there is now an approximately equal number of *cis*-regulatory and coding mutations ($N = 21$ and 20 , respectively). However, the difference between the proportion of *cis*-regulatory mutations

before and after removing the overrepresented genes was not significant (Fisher's exact test, $P = 0.089$).

We performed an additional analysis in which we normalised the data set to weight each individual gene equally regardless of its number of entries in the database (see Section II.6; Fig. S1). The results were again similar, with a slightly increased proportion represented by *cis*-regulatory mutations relative to the unnormalised data set, and association mapping now showing more *cis*-regulatory than coding mutations (59.6%). We conclude that overrepresentation of these melanic case studies may partially explain the higher prevalence of protein coding mutations in the vertebrate pigmentation data set.

(6) The proportion of coding mutations is higher for studies of domesticated and intraspecific variation than for studies of interspecific variation

It has been hypothesised that different mutation types may occur at different frequencies during short-term and long-term evolution, with coding mutations being more common in the short term, and thus across shorter taxonomic distances (Stern & Orgogozo, 2009, 2008). Over shorter periods of time, mutations with deleterious pleiotropic effects might become established if alternative, less-pleiotropic mutations do not appear. By contrast, over longer periods of time, non-optimal mutations will be tested across a variety of environments and there will be more opportunity for adaptive mutations without pleiotropic effects to appear and be selected for. In addition, contexts of artificial selection by humans could facilitate the presence of mutations that would normally be counter-selected in the wild (Cieslak *et al.*, 2011; Courtier-Orgogozo & Martin, 2020; Hanly *et al.*, 2021). To test this in vertebrate pigmentation systems, we plotted the proportion of *cis*-regulatory mutations according to the taxonomic status categories: domesticated, intraspecific, and interspecific. The 'domesticated' category included cases of artificial selection by breeders, with pigment trait variations being directly selected in most cases. The 'intraspecific' category contained studies that investigated natural phenotypic differences between morphs of the same species. 'Interspecific' entries were all those that reported pairs of taxa above the species level.

We found a similarly low proportion of *cis*-regulatory mutations in the domesticated and intraspecific variation categories (17.6 and 24.8%, respectively; Figs 3A and S2). By contrast, studies on interspecific variation reported a significantly higher proportion of *cis*-regulatory mutations (64.3%), suggesting that this mutation type indeed might be more common with increasing taxonomic distance (Fig. 3A). This pattern broadly remained in a re-analysis of the data normalised for number of entries (Fig. S3A), wherein slightly higher proportions of *cis*-regulatory mutations were calculated than in the non-normalised analysis but increasing taxonomic distance was still associated with a higher proportion of *cis*-regulatory mutations.

These results must be interpreted with caution as the number of interspecific studies addressing vertebrate pigmentation was small – only a total of 14 entries were interspecific, corresponding to 11 unique genes. Further, 8 out of

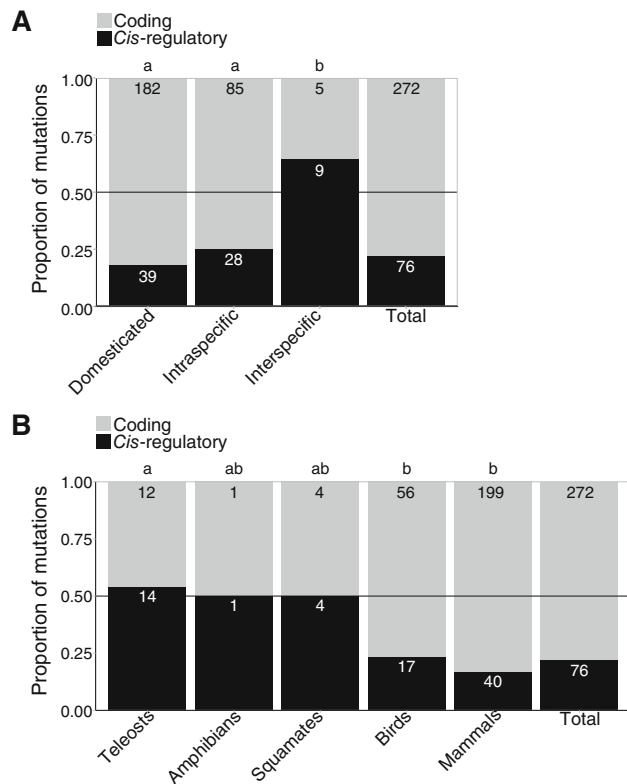


Fig. 3. (A) Proportion of *cis*-regulatory and coding mutations associated with each taxonomic status, and in the total data set. (B) Proportion of *cis*-regulatory and coding mutations associated with each clade. Letters above columns indicate results of *post hoc* pairwise comparisons with Tukey correction ($P < 0.05$); categories that are not significantly different share a letter. Numbers within each bar represent the number of entries in that category.

14 ‘interspecific’ entries were from teleost fishes, which, unlike mammals, have multiple pigment cell types (Fig. S4). The low number of interspecific studies together with their low clade diversity limits the examination of trends across large evolutionary timescales and may reflect an underexplored aspect of vertebrate pigmentation evolution.

(7) The proportion of coding mutations is higher in mammal and bird studies than in studies of teleosts, amphibians, and squamates

We also investigated whether the prevalence of coding mutations could be associated with a disparity in representation of different vertebrate clades. Mammals and birds constitute the vast majority of entries (89.5%, $N = 348$ combined), with entries for teleosts, amphibians, and squamates being much less common (10.5%, $N = 41$ combined). The latter three clades also have much more varied pigmentation systems than mammals or birds, with multiple developmentally distinct classes of pigment cells. This difference could potentially lead to differences in evolutionary trends. For example, the

development of distinct pigment cell classes from a shared neural crest origin increases the likelihood that coding sequence mutations in genes contributing to pigment cell development will have pleiotropic effects. Mutations, and particularly coding mutations, affecting genes with shared roles between chromatophore classes could thus impact several aspects of pigmentation simultaneously. Thus, *cis*-regulatory mutations may perhaps be selected for their specificity of expression as they are more likely to affect only one individual chromatophore class. Similarly, genes implicated in direct cellular interactions might be expected to suffer more negative consequences from coding sequence mutations, given the recognition specificity that is typical of signalling molecules and their receptors. We therefore examined whether clades with distinct pigmentation systems showed different patterns in the proportion of *cis*-regulatory and coding mutations (Fig. 3B). The proportion of *cis*-regulatory changes was lowest for mammals (16.7%) and birds (23.3%). By contrast, studies on squamates (55.6%), teleosts (52.0%) and amphibians (50%; but note that $N = 2$, Fig. 3B) reported approximately equal proportions of the two mutation types.

Although the low number of entries in our data set for taxa with multiple chromatophore pigment systems precludes definitive conclusions, the observed trend does fit with the hypothesis that variation in pigmentation systems utilising multiple pigment cell classes could lead to a higher proportion of *cis*-regulatory mutations. The proportion of *cis*-regulatory mutations for teleost, amphibian, and squamate entries combined was 52.8%, compared with 18.3% for birds and mammals combined. This difference was statistically significant (Fisher’s exact test, $P < 0.0001$). Notably, half of the interspecific data came from teleosts (Fig. S4), making it difficult to draw conclusions regarding patterns in the proportion of *cis*-regulatory mutations across higher taxonomic levels. However, considering only the intraspecific cases, there is a higher proportion of *cis*-regulatory mutations for teleosts, amphibians, and squamates (45.0%, 9/20 entries) than in birds and mammals (20.4%, 19/93) (Fig. S4). This difference was statistically significant (Fisher’s exact test, $P = 0.0419$). Re-analysis of the data using equal weighting for each gene (see Section II.6) for each clade also implies that clades with multiple pigment cell classes exhibit higher proportions of *cis*-regulatory mutations (Fig. S3B). A larger sample of studies conducted in model systems with multiple pigment cell classes, together with more interspecific studies in other clades, would provide valuable insights.

(8) Types of phenotypic variation associated with coding and *cis*-regulatory mechanisms

Differences in the nature of phenotypic variations may be associated with distinct types of genotypic change. For instance, loss of pigmentation may be more likely to result from loss-of-function coding sequence changes, whereas patterning changes could preferentially be driven by *cis*-regulatory mutations due to their high spatial modularity reducing pleiotropic effects (Orteu & Jiggins, 2020).

To determine whether there are differences between colour loss, colour tuning, and pattern variation, we assigned different phenotype categories to each entry (Table 2; see Section II.3).

In every category except changes in carotenoid colouration, coding sequence changes were the most common (Fig. 4A). For the melanic pigmentation categories (pheomelanism, eumelanism, dilution, amelanism, white spotting)

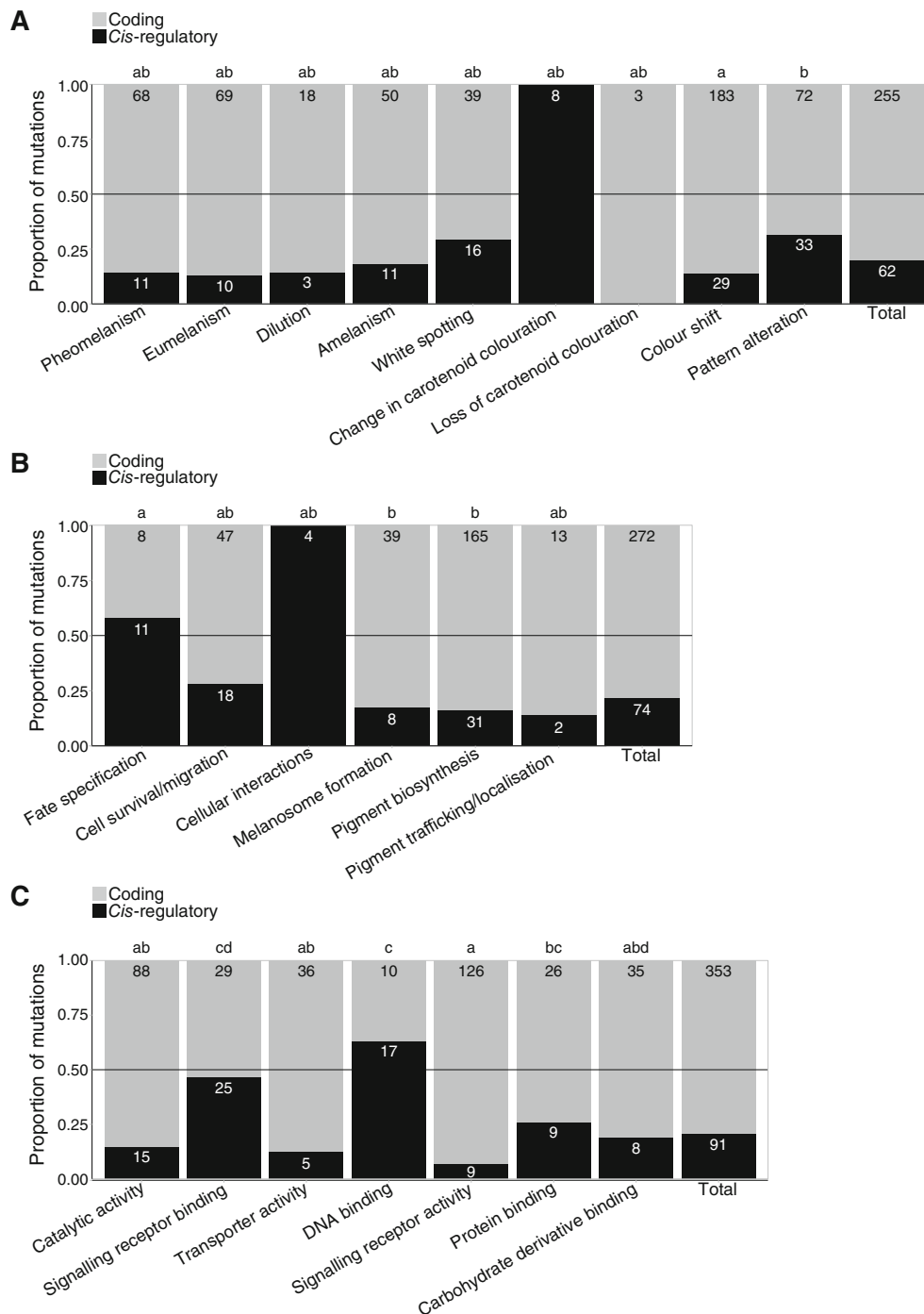


Fig. 4. (A) Relative proportions of *cis*-regulatory and coding mutations associated with each phenotype category, and in the total data set. Note that the total proportions were only for data that could be assigned to either colour shift or pattern alteration. (B) Relative proportions of *cis*-regulatory and coding mutations associated with each cellular function category. (C) The proportion of *cis*-regulatory and coding mutations associated with each molecular function category. Letters above columns indicate results of *post hoc* pairwise comparison with Tukey correction ($P < 0.05$); categories that are not significantly different share a letter. Numbers within each bar represent the number of entries in that category.

cis-regulatory mutations represented a low proportion of the data set (13–29%). These results do not support the suggestion that phenotypes involving loss of pigmentation are associated with a higher proportion of coding mutations than changes to pigmentation quantity or distribution.

Both phenotype categories pertaining to carotenoid colouration had too few entries for reliable conclusions to be drawn. However, both carotenoid categories combined appear to involve a significantly higher proportion of *cis*-regulatory mutations (72.7%) compared with melanic pigmentation (17.3%; Fisher's exact test, $P = 0.0001$), suggesting that the evolution of carotenoid pigmentation may be driven by *cis*-regulatory changes to a greater extent than for melanic pigmentation. The physiological relevance of carotenoids outside of their signalling roles, their distribution across multiple tissue types other than pigment cells, and the plasticity associated with carotenoid pigmentation may mean that carotenoid pigmentation has evolved under different selection pressures compared with melanic pigmentation (Svensson & Wong, 2011). Further studies investigating carotenoid phenotypes would be invaluable to determine whether carotenoid colouration does evolve through different mechanisms compared to melanic pigmentation.

For melanic pigmentation, there were no significant differences between categories. The category with the highest *cis*-regulatory proportion was white spotting (29.1%), followed by amelanism (18.0%), pheomelanism (13.9%), dilution (14.3%), and eumelanism (12.7%). Notably, two phenotype categories were strongly associated with 'domesticated' entries: 18 out of 20 dilution entries, and all 55 entries for white spotting, were identified in domesticated taxa (Fig. S5). White spotting phenotypes are associated with strong pleiotropic effects, with many of these entries in *Gephebase* being deleterious in the homozygous state. It is therefore unsurprising that white spotting mutations were only identified in domesticated taxa, where they can be subject to heterozygous advantage (Hanly *et al.*, 2021; Hedrick, 2015). This may reflect the nature of white spotting phenotypes, which result from impairment of pigment cell migration or differentiation. It is possible that coding mutations that cause white spotting are more likely to be non-viable due to their pleiotropic effects on upstream cellular development.

Phenotypes relating to pigment pattern alteration were associated with a significantly higher proportion of *cis*-regulatory mutations (31.4%) than phenotypes relating to colour shifts (13.7%) (Fisher's exact test, $P = 0.0003$; Fig. 4A). This general pattern was seen for every study methodology and taxonomic status (Figs S5 and S6), although many comparisons did not reach statistical significance. This difference between colour and pattern phenotypes appears to be most pronounced in non-domesticated animals: for intraspecific and interspecific entries combined, pattern phenotypes had a *cis*-regulatory proportion of 63.6% compared with 16.8% for colour (Fisher's exact test, $P < 0.0001$; Fig. S5). This is despite white spotting phenotypes being almost exclusively found in domesticated animals as well as exclusively categorised as pattern alterations. Together, these

results imply that there may be a disparity between colour and pattern phenotypes in terms of the mutation type involved, particularly in the evolution of non-domesticated animals.

Repeating this analysis after correction for the number of entries for each gene (Fig. S7A), again increased the overall proportions represented by *cis*-regulatory mutations for each category. Both eumelanism (46.9%) and amelanism (47.9%) appeared to involve a higher proportion of *cis*-regulatory mutations than other melanic pigmentation categories (pheomelanism 27.5%, dilution 25.3%, white spotting 34.3%), mainly due to the large numbers of *tyr* (amelanism) and *mc1r* (eumelanism) entries in these categories in our data set. Therefore, this comparison also does not support the hypothesis that loss-of-pigmentation phenotypes are more likely to be due to coding mutations. Additionally, this analysis provides further support for a disproportionate role of *cis*-regulatory changes in the generation of pattern variation (53.0%) compared with colour shifts (36.9%).

(9) Upstream developmental processes are associated with a higher proportion of *cis*-regulatory mutations

Differences in mutation types may also be associated with the developmental or cellular function fulfilled by the gene in question. We hypothesised that genes associated with upstream developmental processes would exhibit different proportions of mutation types compared with those associated with downstream processes. For example, genes that play a role in pigment cell fate specification (an upstream process) may be less tolerant of coding sequence mutations than genes contributing only to pigment deposition (a downstream process). To test this, we assigned each gene to one of six categories reflecting different cellular and/or developmental processes (Table 3). Three of these categories are cell type-specific (downstream) functions: melanosome formation, pigment biosynthesis, and pigment trafficking/localisation. The remaining categories involved upstream processes: pigment cell fate specification, pigment cell precursor survival/migration, and cellular interactions (Table 3).

The three categories representing the more upstream cellular/developmental processes appeared to be associated with a higher proportion of *cis*-regulatory mutations than the categories representing downstream cellular functions, although many of these differences were not statistically significant (Fig. 4B). However, grouping the upstream (37.5%, $N = 88$) and downstream (15.9%, $N = 258$) categories, showed a statistically significant difference between them (Fisher's exact test, $P < 0.0001$). Note that there were only four entries for the cellular interactions category, all of which involved *cis*-regulatory mutations (Fig. 4B); this small number makes it difficult to draw conclusions regarding this category. All categories representing downstream processes, i.e. melanosome formation (17.0%), pigment biosynthesis (15.8%) and pigment trafficking/localisation (13.3%), were

associated with a low proportion of *cis*-regulatory mutations. The overall pattern supports the hypothesis that for upstream processes, such as pigment cell fate specification, mutations are more likely to be *cis*-regulatory when compared with downstream cellular processes which do not necessarily affect cell viability, such as variable rates of pigment deposition. This may reflect a lower likelihood of pleiotropic effects associated with downstream processes.

The differences between upstream and downstream processes persist after correction for the number of entries associated with each gene (Fig. S7B), with the exception that a greater proportion of the pigment biosynthesis category involves *cis*-regulatory mutations (50.4%), reflecting a preponderance of some overrepresented genes, such as *mc1r*, *tyr*, and *tyrp*, within this category. Note, however, that with respect to pleiotropy, correcting for the number of entries per gene may lead to misinterpretation. A large number of entries in our data set for a small number of biosynthesis genes could be explained if such genes are less constrained by pleiotropy, rather than simply reflecting a literature bias. Such considerations mean that these results must be interpreted with caution. Nevertheless, our findings may suggest that pigment biosynthesis genes are more functionally constrained than other downstream genes.

(10) Transcription factors are associated with a higher proportion of *cis*-regulatory mutations than other molecular functions

We were also interested in whether the specific molecular activity of a gene would influence the distribution of mutation types. Coding mutations are constrained by the degree of functional conservation required by the protein's molecular function. Transcription factors implicated in multiple regulatory functions may be more constrained than enzymes that catalyse pigment-specific pathways, as even loss-of-function mutations may be tolerated in enzymes that are functionally specialised and non-essential (Pál, Papp & Lercher, 2006). For example, differential expression of the transcription factor *sox10* resulting from *cis*-regulatory mutations underlies changes to melanogenesis in both pigeons and chickens (Domyan *et al.*, 2014; Gunnarsson *et al.*, 2011). The *sox* family of transcription factors, including *sox10*, are highly pleiotropic and essential for neural crest specification, migration and differentiation (Sarkar & Hochedlinger, 2013). No *sox10* coding mutations were identified in this data set. Conversely, multiple presumptive null coding mutations have been identified in *mc1r*, including one example of parallel evolution in the cavefish *Astyanax mexicanus* (Gross, Borowsky & Tabin, 2009). Thus, in contrast to *sox10* and other transcription factors, highly specialised proteins such as *mc1r* may be less susceptible to pleiotropy arising from null mutations. Thus, we predicted that molecular functions that are less pleiotropic would exhibit higher proportions of *cis*-regulatory mutations, and *vice versa*.

Using EBI's QuickGO mouse slimmer and manual combination of highly related molecular GO terms, we assigned each entry to one of seven categories of molecular function

(see Section II.3, Table 4) and examined the proportions of *cis*-regulatory and coding mutations for each category (Fig. 4C). DNA binding had the highest proportion of *cis*-regulatory mutations of any category (63.0%) and was the only category in which they exceeded coding mutations. Signalling receptor binding had the second highest proportion of *cis*-regulatory mutations (46.3%), followed by protein binding (25.7%), and the lowest proportion was identified for signalling receptor activity (6.7%). Taken together, these results suggest that the ability of transcription factors to evolve coding variants that lead to phenotypic changes may indeed be limited by their pleiotropic roles in development, while their non-coding regions offer a more viable path for phenotypic changes.

One potentially interesting result is that we found a significantly higher proportion of *cis*-regulatory mutations for ligands (signalling receptor binding) than for receptors (signalling receptor activity) (Fisher's exact test, $P < 0.0001$). For most signalling receptor genes, the corresponding ligand was also present in the data set in the signalling receptor binding category. For all ligand/receptor pairs, we found a higher proportion of *cis*-regulatory mutations in the ligand than in its receptor, including the receptor/ligand pairing with the most entries in the data set (*mc1r* receptor 1.1%, its ligand *asip* 41.6%). A similar example is provided by *kit*/*kitlg*. For the *kit* receptor, the proportion of *cis*-regulatory mutations was 17.5%, with multiple putative null mutations reported. Although *kitlg* has only four entries in our data set, three are *cis*-regulatory. Both the low number of *kitlg* entries and its relatively high proportion of *cis*-regulatory mutations may indicate lower evolutionary tolerance of coding mutations when compared with its receptor. Overall, our results suggest that the coding sequence of genes encoding signalling ligands is more constrained than that of their signalling receptors, and support previous findings that *cis*-regulatory mutations in ligand genes drive morphological evolution (Martin & Courtier-Orgogozo, 2017).

It has been hypothesised that ligands may play a role in altering spatially localised phenotypes due to the relatively high modularity of their expression patterns (Martin & Courtier-Orgogozo, 2017). For example, in deer mice (*Peromyscus* spp.) alternative isoforms of *asip* are differentially expressed in different body regions, so that *cis*-regulatory mutations can drive spatially localised pattern changes by altering relative isoform expression (Mallarino *et al.*, 2017). Additionally, it has been shown that functional residues within regions of protein interaction are more highly conserved than residues outside of interaction regions, thus it may be the case that only the ligand-binding domains of receptor proteins are highly constrained by specificity, whereas other domains are more tolerant of coding mutations (Worth, Gong & Blundell, 2009). Alternatively, ligand genes, which typically encode a smaller protein, may be expected to have a greater proportion of their coding sequence functionally constrained when compared with the corresponding receptor genes. Taken together, both the modularity associated with ligand expression patterns as well as the greater number of

mutation-tolerant domains in receptors may explain the higher proportion of *cis*-regulatory mutations in ligand-encoding genes compared to receptor genes.

The same general patterns were observed when repeating this analysis after correction for the number of entries associated with each gene (Fig. S7C), although absolute proportions of *cis*-regulatory mutations were higher for most categories. The category with the largest increase compared with the uncorrected data (Fig. 4C) was catalytic activity, primarily because of the large number of coding entries in this category for *tyr* and *tyrp*.

(11) Divergent pigmentation systems have fewer shared genetic hotspots

Studies of the genetic basis of evolutionary change have shown that repeated co-option of the same genes underlies the evolution of convergent phenotypes. Genes that are repeatedly found as causal drivers of similar phenotypic changes in diverse taxa or populations have been referred to as ‘genetic hotspots’ (Martin & Orgogozo, 2013). We hypothesised that similar pigmentation systems would be more likely to evolve through the same sets of genes, whereas more divergent pigmentation systems would have fewer shared genetic hotspots.

In our comparison of normalised gene overlap between clades (see Section II.4), the greatest overlap was found for mammalian entries for genes shared with birds, and in bird entries for genes shared with mammals (Fig. 5). For squamates, the greatest overlap was for genes shared with teleosts. Teleosts had comparatively few shared hotspot genes with any other clade, with the greatest overlap being with mammals, possibly due to the large number of teleost entries related to melanic phenotypes (65.5%, 19 out of 29 entries). Additionally, there was considerable overlap between birds and squamates which may reflect their phylogenetic closeness. As the number of

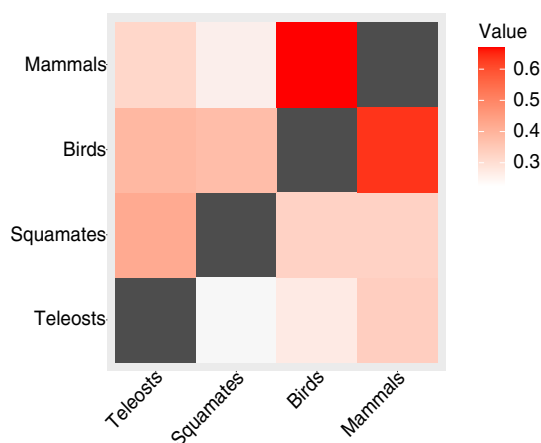


Fig. 5. Heatmap of the normalised gene overlap in our data set between each clade pairing. For example, the top left square has a value equal to the number of mammalian entries for genes that were also identified in teleost studies, divided by the total number of mammalian entries.

shared hotspot genes would be expected to reduce over evolutionary time, phylogenetic distance is likely to contribute to the observed overlap between clades (Conte *et al.*, 2012).

Although *mc1r* was highly represented in the data set for all clades, and thus contributed disproportionately to our hotspot calculations for all clade pairings, many other genes that were highly represented in the data set were confined to mammals and birds only. The second, third, and fourth most represented genes, *kit*, *asip*, and *tyrp1*, which all function in development of melanic pigments, had no teleost or squamate entries. However, other melanogenic genes exhibited overlap between teleost and squamate entries, such as *oca2*, a melanosomal transport protein (Table S1). In total, *oca2* comprised 5 out of 29 (17.2%) teleost entries, and 1 out of 10 (10%) squamate entries, two in mammals (0.82%) and was absent entirely in birds. Only one non-melanic pigmentation gene was shared between birds and squamates. Aside from *mc1r*, no single gene was identified in all four clades.

Overall, the observed trend of higher hotspot overlap between similar pigmentation systems may indicate shared genetic mechanisms underlying pigment evolution in systems with or without multiple pigment cell classes. The low number of non-melanic pigmentation entries (only 31 in total) makes it difficult to draw conclusions about the relevance of hotspot genes outside of melanogenic pathways for clades with additional pigment types.

IV. DISCUSSION

Gephebase is a powerful tool that provides an accessible compilation of data on phenotype–genotype associations to allow investigations of general patterns regarding the mechanisms underlying evolutionary change. Importantly, this data set can be used to identify gaps in the literature, and to identify novel directions and study systems for future research. Here, we analysed a large data set of 389 *Gephebase* entries, each representing a pair of alleles linked to phenotypic variation in vertebrate pigmentation. Of the three forms of empirical evidence accepted by *Gephebase*, it was clear that the study methodology with the highest ascertainment bias – the candidate gene approach – identified more coding sequence mutations than the other methods. Irrespective of this bias, protein coding changes were still the majority of mutations identified in all analyses, representing 75% of entries for linkage mapping studies and 50% for association mapping studies. In addition, we identified several factors that may affect the proportions of coding and *cis*-regulatory mutations. Below we discuss these factors and highlight new avenues of research in the vertebrate pigmentation field.

(1) Strong selection pressure might explain the higher proportion of coding mutations associated with vertebrate pigmentation diversity

A majority of studies on vertebrate pigmentation, particularly in domesticated species, identified coding mutations.

In the case of domesticated vertebrates, this may be explained by the strong selective pressures inherent in the domestication process. During any form of artificial selection, desirable traits will be under strong directional selection, potentially outweighing any negative pleiotropic effects of coding mutations. One example is that of ‘frame overo’ horses, where changes in patterning are a result of heterozygosity for a coding mutation in the endothelin B receptor (*ednrb*), which leads to a white spotted pattern (Metallinos, Bowling & Rine, 1998). Homozygosity for this mutation leads to lethal white foal syndrome – affected foals are completely white and die shortly after birth. Selection for white spotted patterns in domestic horses has thus led to a lethal deleterious coding mutation becoming established, whereas in wild populations such a mutation would likely not have conferred sufficient advantage to overcome its negative pleiotropy.

A sufficiently strong selection pressure may likewise allow the emergence of pleiotropic coding mutations in wild populations. *Gephebase* contains examples that demonstrate how rapidly changes in pigmentation patterns may evolve when under such pressures, particularly in the context of cryptic camouflage in a new or rapidly changing environment. Populations of North American deer mice exhibit high gene flow, and thus low population genetic structure, with the exception of the agouti (*asip*) locus. Strong selection for cryptic against variable soil colouration has led to high variance in the *asip* coding sequence, and rapid divergence driven by colonisation over the last ~4000 years (Pfeifer *et al.*, 2018). The numerous intraspecific coding mutations in our data set may therefore be indicative of the strength of selection typical for vertebrate pigmentation evolution. In the future it would be interesting to investigate the relative strength of selection on each mutation type, and to perform the same analyses for multiple morphological traits to determine if the observed trends are pigmentation specific.

(2) Evolutionary timescales and taxonomic distance

We found that shorter taxonomic distances (i.e. domesticated and intraspecific entries) were associated with a higher proportion of protein coding changes than in interspecific entries. These results should be taken with caution as the smaller number of interspecific studies indicates comparatively little investigation of pigmentation traits across larger evolutionary timescales. Interestingly, the majority (8 out of 14) of interspecific entries in our data set were for teleosts, despite their overall underrepresentation. Four of these were from East African cichlids, a group of teleost fishes characterised by rapid diversification and speciation. The observed higher frequency of *cis*-regulatory mutations found in interspecific studies thus could be biased by a focus on these particular cases of rapid evolution.

The current overrepresentation in the data set of shorter evolutionary timescales may itself introduce bias in the types of mutation that can be identified. Shorter evolutionary timescales may allow for pleiotropic mutations, as well as

mutations that confer an advantage only in a specific environmental context. Thus, the higher number of intraspecific and domesticated case studies is likely to result in overrepresentation of coding mutations. Expanding the timescales investigated in pigmentation evolution will facilitate testing of this hypothesis, and a greater understanding of how pigmentation changes over evolutionary time.

Identifying bird and mammal model systems for interspecific studies is challenging owing to their often-limited potential for hybridisation, restricting the types of mapping techniques available. For instance, the only two interspecific mammalian entries examined differences within the black rat (*Rattus spp.*) species complex, and between the house mouse (*Mus musculus*) and tobacco mouse (*Mus poschiavinus*), respectively (Kambe *et al.*, 2011; Robbins *et al.*, 1993). Both of these studies used a candidate gene approach. In birds, Toews *et al.* (2016) used an association mapping approach to examine pigmentation evolution in two naturally hybridising species of warbler with highly divergent pigment patterns. Diversification of model systems in the future may facilitate a greater understanding of long-term pigment evolution in birds and mammals.

(3) Potential biases introduced by study systems

Mammalian and avian systems dominate our data set. These taxa possess only melanocytes as pigment cells, and pigmentation differences primarily result from spatial and temporal differences in pigment production and deposition. In other vertebrates, pigmentation patterns result from the presence and differential arrangement of different classes of pigment cells that share a developmental origin. The few studies in the *Gephebase* database that focus on clades with distinct chromatophore classes often found novel pigmentation phenotypes associated with a wider range of genes. This is corroborated by the greater numbers of shared hotspot genes between clades with similar pigmentation systems – there was considerable overlap between mammal and bird hotspot genes. Studies in clades with distinct chromatophore classes also implicated a greater role for *cis*-regulatory mutations. As current research continues to contribute new data, it will be interesting to investigate whether this trend is confirmed, or whether the observed higher prevalence of *cis*-regulatory mutations in teleosts, squamates and amphibians is an artefact of the small number of studies available to date. A greater emphasis on a broader set of model systems will enhance our understanding of vertebrate pigmentation evolution.

It would be of interest to expand the number of studies on teleosts and reptiles to validate the hypothesis that pigmentation systems with multiple chromatophore classes evolve preferentially through *cis*-regulatory modifications. Similarly, amphibian pigmentation evolution was only represented by two studies in our data set. Amphibians’ highly permeable skin and lack of epidermal protection, as well as their commonly biphasic aquatic and terrestrial life cycles, may necessitate distinct adaptive functions for colouration compared with other vertebrate clades (Rudh &

Qvarnström, 2013). Additional information for these under-represented clades will be highly informative, particularly with respect to interactions between chromatophore classes and pattern formation.

(4) The importance of understanding the cellular and developmental basis of variation

We found a higher proportion of *cis*-regulatory mutations for changes in patterning than for changes in colour. This is consistent with previous hypotheses that *cis*-regulatory mutations largely govern spatial changes due to altering expression patterns (Carroll, 2008). For example, in sticklebacks (*Gasterosteus aculeatus*) a *cis*-regulatory allele of the kit ligand gene reduces its specific expression in gill tissue, leading to divergent gill pigmentation phenotypes (Miller *et al.*, 2007). *Gephebase* provides several such examples, where mutations in widely expressed genes are *cis*-regulatory, so that expression only in specific tissues is affected.

Nonetheless, a majority of pattern alteration case studies still involve coding mutations. When looking more closely at these, we found that many involve loss-of-function mutations in genes that are expressed only in specific regions of the body. For instance, the urucum breed of canary (*Serinus canaria*) exhibits a red bill and legs, which are normally yellow in wild-type canaries (Gazda *et al.*, 2020). This is the result of a mutation that compromises the enzymatic activity of beta-carotene oxidase 2 (*bco2*), an enzyme responsible for breakdown of full-length red carotenoids into shorter apocarotenoids. The specificity of this allele in affecting the bill and the legs is the result of loss-of-function mutations in a gene whose expression is restricted to specific tissues. Without knowing the expression patterns and the level of pleiotropy of a given gene, we may be wrongly categorising phenotypes as pattern alterations when in fact they are colour shifts. Unfortunately, we have no way of testing this, since most case studies focus solely on the genotype–phenotype association without conducting gene expression or developmental studies. Knowledge of a gene's specific function, together with its placement in a cellular and developmental context, is essential for an understanding of when and how certain genes and mutations are favoured over others.

Characterising trait development at the gene-expression and cellular level is also important since we found potential differences in mutation proportions associated with the upstream or downstream position of cellular and developmental processes. The genes contributing to upstream processes, i.e. those related to fate specification and cellular interactions, were more likely to be *cis*-regulatory compared with downstream processes. A previous study (Stern & Orgogozo, 2008) distinguished two categories of genes: those that act at or near the terminal points of regulatory networks, named 'differentiation gene batteries', and those that are upstream in the gene network. They found that the proportion of *cis*-regulatory mutations was not significantly different for the two categories. This analysis was based on a compilation of 331 mutations for all types of traits across animals and

plants. It is possible that this previous categorization was affected by the very diverse gene networks considered. Here, by focusing only on vertebrate pigmentation, we uncovered a potential effect of upstream *versus* downstream position on the proportion of *cis*-regulatory mutations. While these results contradict those of Stern & Orgogozo (2008), this may reflect an effect specific to vertebrate pigmentation. Additional studies, particularly focusing on upstream genes, may provide useful clarification.

One study identified a *cis*-regulatory mutation that altered the expression of colony stimulating factor 1 (*csf1*), a signalling factor expressed in xanthophores that is required for their differentiation and survival (Patterson, Bain & Parichy, 2014). In *Danio albolineatus*, a close relative of the zebrafish *D. rerio*, increased expression *via* a *cis*-regulatory mutation of *csf1a* was associated with early xanthophore recruitment (Patterson *et al.*, 2014). This precocious xanthophore appearance led to changes in abundance and distribution of all three principal chromatophore classes (melanophores, xanthophores, and iridophores) which ultimately inhibited formation of the stripe pattern found in other *Danio* species. This study illustrates the potential complexity of phenotypes arising from mutations that affect cellular interactions. The identification of such mutations represents an exciting avenue for future pigmentation evolution research.

By contrast, coding mutations in downstream processes would typically be expected to have less impact on non-pigmentation phenotypes. One study found that a loss-of-function mutation in scavenger receptor B1 (*scarb1*) in canaries produces a completely white phenotype due to *scarb1* being necessary for carotenoid uptake. As a result of the specialised role of *scarb1*, this null coding mutation produces a spectacular pigmentation phenotype (Toomey *et al.*, 2017). Together, these results support the hypothesis that developmentally upstream genes may be more constrained by pleiotropy. Given the over-representation of pigment synthesis case studies in the vertebrate pigmentation data set, the collection of additional data regarding other types of trait variation is imperative.

(5) What are the main knowledge gaps?

Most vertebrate pigmentation entries (363 out of 389) in our data set result from the study of variation in melanic traits. By contrast, variation in other types of colour traits, such as carotenoids and pteridines, structural colouration, and colour plasticity remain largely unexplored. Further, most case studies address naturally selected traits with a relatively simple genetic basis – only 21 out of 389 entries assessed variable sexually dimorphic traits (see supporting information, File S1, tab: All Vertebrates). It would therefore be interesting to investigate the proportions of mutation types for sexually selected traits or highly polygenic traits where each mutation potentially has a small effect.

A focus on previously unexplored traits may allow identification of uncharacterised genes underlying these traits. For example, the study of variation in yellow pigmentation in budgerigars (*Melopsittacus undulatus*) led to the identification of a coding mutation in a previously uncharacterised

polyketide synthase that is involved in the synthesis of yellow polyene pigments called psittacofulvins (Cooke *et al.*, 2017).

We found only four entries in the data set relating to structural colouration, all of which involved iridophore distribution. Three of these entries involved phenotypes affecting multiple chromatophore classes, of which two were classified as belonging to the ‘cellular interactions’ category (Table 3). For example, in an interspecific study, Spiewak *et al.* (2018) investigated pigment pattern differences between two *Danio* fish species. *D. nigrofasciatus* has a disrupted stripe pattern compared to *D. rerio*, which is associated with a *cis*-regulatory mutation in the endothelin-3 (*edn3*) signalling peptide (Spiewak *et al.*, 2018). Because *edn3* promotes iridophore proliferation, this *cis*-mutation leads to decreased *edn3* expression and correspondingly lower iridophore complement in *D. nigrofasciatus*. A lower number of iridophores secondarily affects melanophore proliferation, leading to a reduction of both cell types and disruption of the defined striped pattern in *D. nigrofasciatus*. The fourth structural colouration entry involved changes only to iridophores. This study investigated ectopic iridophore proliferation in a domesticated leopard gecko morph, implicating a serine peptidase inhibitor, *spint1*, in iridophore abundance and the development of skin tumours. This entry was the only study for *spint1*, and its specific role in iridophore biology remains unclear. The low number of *Gephebase* entries relating to structural variation highlights that structural colouration is relatively unexplored, and represents an emerging field of research where new genes and developmental processes leading to variation in cells and physical structures will be identified.

The current entries in *Gephebase* are heavily focused on colouration traits that are genetically determined independently of environmental variables. Colour variation can also arise as a result of phenotypic plasticity, where the same genotype can generate different colour states under different conditions. Six entries pertained to plastic or seasonal colour changes – for example, *cis*-regulatory alleles of *asip* determine whether the winter coat of snowshoe hares is brown or white (Jones *et al.*, 2018). Determining how plasticity emerges and evolves remains a challenge. Genes involved in plasticity evolution can only be identified when mapping differences between closely related species or populations that show variation in the presence and absence of plastic colouration (Gibert, 2017). Likewise, variation in plastic colour responses can only be mapped by measuring variation in population reaction norms, which are costly and time-consuming experiments. Nonetheless, such studies will be key to differentiate between the two main hypotheses regarding the evolution of plasticity: whether evolutionary change occurs through changes in genes that sense and regulate a downstream response to external factors, or instead through changes in the colouration genes themselves that become responsive to environmental cues.

Sexually selected traits are also underrepresented in the database, making up only 5.4% of the case studies. Understanding the interplay between natural and sexual selection on colour traits is important since they may often act in

opposite directions. For example, in several species of Lake Malawi cichlids a well-camouflaged colour morph is associated with a *cis*-regulatory mutation in *pax7*. However, this mutation also has a deleterious effect in that it disrupts male nuptial colouration. This conflict has been resolved by the invasion of a novel sex determination locus in tight linkage with the *pax7* allele (Roberts, Ser & Kocher, 2009). The importance of pigmentation patterns in mate choice can lead to such conflicts, and the ways in which they are resolved could provide fascinating case studies. Sexually selected traits also present a conundrum in that it is unclear why and how trait variation is maintained in natural populations despite apparent directional selection due to female choice or male–male competition (Chenoweth & McGuigan, 2010). Identifying genes and genomic regions contributing to trait variation and studying its adaptive significance in wild populations creates the opportunity to understand this paradox (Johnston *et al.*, 2013).

Finally, the *Gephebase* data set is currently biased towards large-effect loci, as linkage mapping is limited in its statistical power to detect small-effect mutations and polygenic architectures (Rockman, 2012; Slate, 2013). Association studies conducted on thousands of samples are now reaching a point where small-effect loci are detectable, under the conditions that these variants are common: for instance, several GWAS studies of human skin pigmentation levels have uncovered an amino acid polymorphism of *mfsd12* as a determinant of colour variation across several continents (Adhikari *et al.*, 2019; Crawford *et al.*, 2017; Feng, McQuillan & Tishkoff, 2021; Lona-Durazo *et al.*, 2019). This gene is now a candidate that is showing association signals in domestic animal studies (Hédan *et al.*, 2019; Tanaka *et al.*, 2019), suggesting it is effectively a hotspot of pigment variation. While human GWAS studies are systematically curated and will undoubtedly lead to powerful meta-analyses (Buniello *et al.*, 2019), and while infrastructure is being created to integrate these data with laboratory model systems (Shefchek *et al.*, 2020), there is a lack of resources to compile data from evolutionary and artificial breeding gene-to-trait relationships beyond these organisms. *Gephebase*, *OMIA*, and *Animal QTLdb* (Courtier-Orgogozo *et al.*, 2020; Hu, Park & Reecy, 2021; Lenffer *et al.*, 2006) represent current attempts to curate these data, but long-term resources are needed to keep pace with increased rates of discovery in the genomic age.

V. CONCLUSIONS

(1) As more genetic variants underlying trait variation are identified, it becomes possible to test more rigorously predictions relating to the mechanisms of genetic evolution. Here, we highlight some of the trends in studies of vertebrate pigmentation evolution, and specifically test some of the predictions made about the relative frequencies of *cis*-regulatory and coding mutations. In contrast to expectations, we found that the majority of the documented variation in pigmentation is driven by coding sequence mutations.

(2) However, we also identified multiple factors potentially associated with the observed proportions that partly explain this disparity. We make suggestions for the future direction of vertebrate pigmentation research with respect to both systems and study design. As the number and variety of case studies continues to increase, we expect our understanding of the genetic, cellular and developmental mechanisms underlying the evolution of vertebrate pigmentation to expand.

VI. ACKNOWLEDGEMENTS

We thank members of the Morphological Evolution Lab for discussion. A. M. and V. C.-O. were supported by a John Templeton Foundation grant from 2014 to 2017 (JTF 43903). M. E. S. is supported by NERC IRF NE/R01504X/1.

VII. REFERENCES

References identified with an asterisk (*) are cited in File S1 included as online supporting information.

- *ABITBOL, M., BOSSÉ, P., GRIMARD, B., MARTIGNAT, L. & TIRET, L. (2017). Allelic heterogeneity of albinism in the domestic cat. *Animal Genetics* **48**, 127–128.
- *ABITBOL, M., LEGRAND, R. & TIRET, L. (2014). A missense mutation in melanocortin 1 receptor is associated with the red coat colour in donkeys. *Animal Genetics* **45**, 878–880.
- *ABITBOL, M., LEGRAND, R. & TIRET, L. (2015). A missense mutation in the agouti signaling protein gene (ASIP) is associated with the no light points coat phenotype in donkeys. *Genetics, Selection, Evolution* **47**, 28.
- *ADHIKARI, K., MENDOZA-REVILLA, J., SOHAIL, A., FUENTES-GUAJARDI, M., LAMPERT, J., CHACÓN-DUQUE, J. C., HURTADO, M., VILLEGAS, V., GRANJA, V., ACUÑA-ALONZO, V., JARAMILLO, C., ARIAS, W., LOZANO, R. B., EVERARDO, P., GÓMEZ-VALDÉS, J., *ET AL.* (2019). A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia. *Nature Communications* **10**, 358.
- *AIGNER, B., BESENFELDER, U., MÜLLER, M. & BREM, G. (2000). Tyrosinase gene variants in different rabbit strains. *Mammalian Genome* **11**, 700–702.
- *ALBERTSON, R. C., POWDER, K. E., HU, Y., COYLE, K. P., ROBERTS, R. B. & PARSONS, K. J. (2014). Genetic basis of continuous variation in the levels and modular inheritance of pigmentation in cichlid fishes. *Molecular Ecology* **23**, 5135–5150.
- *ALMATHEN, F., ELBIR, H., BAHBAHANI, H., MWACHARO, J. & HANOTTE, O. (2018). Polymorphisms in MC1R and ASIP genes are associated with coat color variation in the Arabian camel. *The Journal of Heredity* **109**, 700–706.
- *ANDERSON, T. M., VONHOLDT, B. M., CANDILLE, S. I., MUSIANI, M., GRECO, C., STAHLER, D. R., SMITH, D. W., PADHUKASAHASRAM, B., RANDI, E., LEONARD, J. A., BUSTAMANTE, C. D., OSTRANDER, E. A., TANG, H., WAYNE, R. K. & BARSH, G. S. (2009). Molecular and evolutionary history of melanism in North American gray wolves. *Science* **323**, 1339–1343.
- *ANDRADE, P., PINHO, C., LANUZA, G. P. I. D., AFONSO, S., BREJCHA, J., RUBIN, C.-J., WALLERMAN, O., PEREIRA, P., SABATINO, S. J., BELLATI, A., PELLITTERI-ROSA, D., BOSAKOVA, Z., BUNIKIS, I., CARRETERO, M. A., FEINER, N., *ET AL.* (2019). Regulatory changes in pterin and carotenoid genes underlie balanced color polymorphisms in the wall lizard. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 5633–5642.
- *ANISTOROAEI, R., FREDHOLM, M., CHRISTENSEN, K. & LEEB, T. (2008). Albinism in the American mink (*Neovison vison*) is associated with a tyrosinase nonsense mutation. *Animal Genetics* **39**, 645–648.
- *ANSAI, S., MOCHIDA, K., FUJIMOTO, S., MOKODONGAN, D. F., SUMARTO, B. K. A., MASENGI, K. W. A., HADIATY, R. K., NAGANO, A. J., TOYODA, A., NARUSE, K., YAMAHIRA, K. & KITANO, J. (2021). Genome editing reveals fitness effects of a gene for sexual dichromatism in Sulawesian fishes. *Nature Communications* **12**, 1350.
- *ATTIE, A. D., HAMON, Y., BROOKS-WILSON, A. R., GRAY-KELLER, M. P., MACDONALD, M. L. E., RIGOT, V., TEBON, A., ZHANG, L.-H.,

- MULLIGAN, J. D., SINGARAJA, R. R., BITGOOD, J. J., COOK, M. E., KASTELEIN, J. J. P., CHIMINI, G. & HAYDEN, M. R. (2002). Identification and functional analysis of a naturally occurring E39K mutation in the ABCA1 gene of the WHAM chicken. *Journal of Lipid Research* **43**, 1610–1617.
- *AWASTHI MISHRA, N., DRÖGEMÜLLER, C., JAGANNATHAN, V., KELLER, I., WÜTHRICH, D., BRUGGMANN, R., BECK, J., SCHÜTZ, E., BREINIG, B., DEMMEL, S., MOSER, S., SIGNER-HASLER, H., PIEŃKOWSKA-SCHELLING, A., SCHELLING, C., SANDE, M., *ET AL.* (2017). A structural variant in the 5'-flanking region of the TWIST2 gene affects melanocyte development in belted cattle. *PLoS One* **12**, e0180170.
- *BAIAO, P. C. & PARKER, P. G. (2012). Evolution of the melanocortin-1 receptor (MC1R) in boobies and gannets (Aves, Suliformes). *The Journal of Heredity* **103**, 322–329.
- *BAIAO, P. C., SCHREIBER, E. & PARKER, P. G. (2007). The genetic basis of the plumage polymorphism in red-footed boobies (*Sula sula*): a melanocortin-1 receptor (MC1R) analysis. *The Journal of Heredity* **98**, 287–292.
- *BAIZ, M. D., WOOD, A. W., BRELSFORD, A., LOVETTE, I. J. & TOEWS, D. P. L. (2021). Pigmentation genes show evidence of repeated divergence and multiple bouts of introgression in *Setophaga* warblers. *Current Biology* **31**, 643–649.e3.
- BARRETT, R. D. H. & HOEKSTRA, H. E. (2011). Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics* **12**, 767–780.
- *BECKER, D., OTTO, M., AMMANN, P., KELLER, I., DRÖGEMÜLLER, C. & LEEB, T. (2015). The brown coat colour of Coppernecked goats is associated with a non-synonymous variant at the TYRP1 locus on chromosome 8. *Animal Genetics* **46**, 50–54.
- BEDFORD, N. L. & HOEKSTRA, H. E. (2015). *Peromyscus* mice as a model for studying natural variation. *eLife* **4**, e06813.
- *BELLONE, R. R., FORSYTH, G., LEEB, T., ARCHER, S., SIGURDSSON, S., IMSLAND, F., MAUCELI, E., ENGENSTEINER, M., BAILEY, E., SANDMEYER, L., GRAHN, B., LINDBLAD-TOH, K. & WADE, C. M. (2010). Fine-mapping and mutation analysis of TRPM1: a candidate gene for leopard complex (LP) spotting and congenital stationary night blindness in horses. *Briefings in Functional Genomics* **9**, 193–207.
- *BENKEL, B. F., ROUVINEN-WATT, K., FARID, H. & ANISTOROAEI, R. (2009). Molecular characterization of the Himalayan mink. *Mammalian Genome* **20**, 256–259.
- *BERRYERE, T. G., KERNS, J. A., BARSH, G. S. & SCHMUTZ, S. M. (2005). Association of an Agouti allele with fawn or sable coat color in domestic dogs. *Mammalian Genome* **16**, 262–272.
- *BERRYERE, T. G., SCHMUTZ, S. M., SCHIMPF, R. J., COWAN, C. M. & POTTER, J. (2003). TYRP1 is associated with dun coat color in Dexter cattle or how now brown cow? *Animal Genetics* **34**, 169–175.
- *BLASZCZYK, W. M., DISTLER, C., DEKOMIEN, G., ARNING, L., HOFFMANN, K.-P. & EPPLEN, J. T. (2007). Identification of a tyrosinase (TYR) exon 4 deletion in albino ferrets (*Mustela putorius furo*). *Animal Genetics* **38**, 421–423.
- *BONILLA, C., BOXILL, L.-A., DONALD, S. A. M., WILLIAMS, T., SYLVESTER, N., PARRA, E. J., DIOS, S., NORTON, H. L., SHRIVER, M. D. & KITTLES, R. A. (2005). The 8818G allele of the agouti signaling protein (ASIP) gene is ancestral and is associated with darker skin color in African Americans. *Human Genetics* **116**, 402–406.
- *BOX, N. F., WYETH, J. R., O'GORMAN, L. E., MARTIN, N. G. & STURM, R. A. (1997). Characterization of melanocyte stimulating hormone receptor variant alleles in twins with red hair. *Human Molecular Genetics* **6**, 1891–1897.
- *BREINIG, B., BECK, J., FLOREN, C., BORNEMANN-KOLATZKI, K., WIEDEMANN, I., HENNECKE, S., SWALVE, H. & SCHÜTZ, E. (2013). Molecular genetics of coat colour variations in white Galloway and White Park cattle. *Animal Genetics* **44**, 450–453.
- *BROOKS, S. A. & BAILEY, E. (2005). Exon skipping in the KIT gene causes a Sabino spotting pattern in horses. *Mammalian Genome* **16**, 893–902.
- *BRUDERS, R., VAN HOLLEBEKE, H., OSBORNE, E. J., KRONENBERG, Z., MACLARY, E., YANDELL, M. & SHAPIRO, M. D. (2020). A copy number variant is associated with a spectrum of pigmentation patterns in the rock pigeon (*Columba livia*). *PLoS Genetics* **16**, e1008274.
- *BRUNBERG, E., ANDERSSON, L., COTHRAN, G., SANDBERG, K., MIKKO, S. & LINDGREN, G. (2006). A missense mutation in PMEL17 is associated with the Silver coat color in the horse. *BMC Genetics* **7**, 46.
- BUNIELLO, A., MACARTHUR, J. A. L., CEREZO, M., HARRIS, L. W., HAYHURST, J., MALANGONE, C., MCMAHON, A., MORALES, J., MOUNTJOY, E., SOLLIS, E., SUVEGES, D., VROUSGOU, O., WHETZEL, P. L., AMODE, R., GUILLEN, J. A., *ET AL.* (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research* **47**, D1005–D1012.
- *CADUFF, M., BAUER, A., JAGANNATHAN, V. & LEEB, T. (2017). A single base deletion in the SLC45A2 gene in a Bullmastiff with oculocutaneous albinism. *Animal Genetics* **48**, 619–621.
- *CANDILLE, S. I., KAELEN, C. B., CATTANACH, B. M., YU, B., THOMPSON, D. A., NIX, M. A., KERNS, J. A., SCHMUTZ, S. M., MILLHAUSER, G. L. & BARSH, G. S. (2007). A β -defensin mutation causes black coat color in domestic dogs. *Science* **318**, 1418–1423.

- *CAPOMACCIO, S., MILANESI, M., NOCELLI, C., GIONTELLA, A., VERINI-SUPPLIZI, A., BRANCA, M., SILVESTRELLI, M. & CAPPELLI, K. (2017). Splicing site disruption in the KIT gene as strong candidate for white dominant phenotype in an Italian Trotter. *Animal Genetics* **48**, 727–728.
- CARROLL, S. B. (2008). Evo-Devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36.
- *CHANDRAMOHAN, B., RENIERI, C., LA MANNA, V. & LA TERZA, A. (2013). The alpaca agouti gene: genomic locus, transcripts and causative mutations of eumelanic and pheomelanic coat color. *Gene* **521**, 303–310.
- CHENOWETH, S. F. & MCGUIGAN, K. (2010). The genetic basis of sexually selected variation. *Annual Review of Ecology, Evolution, and Systematics* **41**, 81–101.
- *CHANG, C.-M., COVILLE, J.-L., COQUERELLE, G., GOURICHON, D., OULMOUDEN, A. & TIXIER-BOICHARD, M. (2006). Complete association between a retroviral insertion in the tyrosinase gene and the recessive white mutation in chickens. *BMC Genomics* **7**, 19.
- *CHO, Y. S., HU, L., HOU, H., LEE, H., XU, J., KWON, S., OH, S., KIM, H.-M., JHO, S., KIM, S., SHIN, Y.-A., KIM, B. C., KIM, H., KIM, C., LUO, S.-J., ET AL. (2013). The tiger genome and comparative analysis with lion and snow leopard genomes. *Nature Communications* **4**, 2433.
- CIESLAK, M., REISSMANN, M., HOFREITER, M. & LUDWIG, A. (2011). Colours of domestication. *Biological Reviews* **86**, 885–899.
- *CIRERA, S., MARKAKIS, M. N., CHRISTENSEN, K. & ANISTOROAEI, R. (2013). New insights into the melanophilin (MLPH) gene controlling coat color phenotypes in American mink. *Gene* **527**, 48–54.
- *CIRERA, S., MARKAKIS, M. N., KRISTIANSEN, T., VISSENBERG, K., FREDHOLM, M., CHRISTENSEN, K. & ANISTOROAEI, R. (2016). A large insertion in intron 2 of the TYRPI gene associated with American Palomino phenotype in American mink. *Mammalian Genome* **27**, 135–143.
- *CLARK, L. A., TSAI, K. L., STARR, A. N., NOWEND, K. L. & MURPHY, K. E. (2011). A missense mutation in the 20S proteasome $\beta 2$ subunit of Great Danes having harlequin coat patterning. *Genomics* **97**, 244–248.
- *CLARK, L. A., WAHL, J. M., REES, C. A. & MURPHY, K. E. (2006). Retrotransposon insertion in SILV is responsible for merle patterning of the domestic dog. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 1376–1381.
- *CONE, R. D., LU, D., KOPPULA, S., VAGE, D. I., KLUNGLAND, H., BOSTON, B., CHEN, W., ORTH, D. N., POUTON, C. & KESTERSON, R. A. (1996). The melanocortin receptors: agonists, antagonists, and the hormonal control of pigmentation. *Recent Progress in Hormone Research* **51**, 287–317 discussion 318.
- CONTE, G. L., ARNEGARD, M. E., PEICHEL, C. L. & SCHLUTER, D. (2012). The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B: Biological Sciences* **279**, 5039–5047.
- *COOK, D., BROOKS, S., BELLONE, R. & BAILEY, E. (2008). Missense mutation in exon 2 of SLC36A1 responsible for champagne dilution in horses. *PLoS Genetics* **4**, e1000195.
- *COOKE, T. F., FISCHER, C. R., WU, P., JIANG, T.-X., XIE, K. T., KUO, J., DOCTOROV, E., ZEHNDER, A., KHOSLA, C., CHUONG, C.-M. & BUSTAMANTE, C. D. (2017). Genetic mapping and biochemical basis of yellow feather pigmentation in budgerigars. *Cell* **171**, 427–439.e21.
- *CORI, A., BI, K., LUKE, C., CHALLA, A. S., STERN, A. J., SINERVO, B. & NIELSEN, R. (2018). The genetic basis of adaptation following plastic changes in coloration in a novel environment. *Current Biology* **28**, 2970–2977.e7.
- *CORSO, J., MUNDT, N. I., FAGUNDES, N. J. R. & DE FREITAS, T. R. O. (2016). Evolution of dark colour in toucans (Ramphastidae): a case of molecular adaptation? *Journal of Evolutionary Biology* **29**, 2530–2538.
- *CORTIMIGLIA, C., CASTIGLIONI, B., PIZZI, F., STELLA, A. & CAPRA, E. (2017). Involvement of tyrosinase-related protein 1 gene in the light brown plumage phenotype of *Falco cherrug*. *Animal Genetics* **48**, 125–126.
- COURTIER-ORGOGOZO, V., ARNOULT, L., PRIGENT, S. R., WILTGEN, S. & MARTIN, A. (2020). Gephabase, a database of genotype–phenotype relationships for natural and domesticated variation in Eukaryotes. *Nucleic Acids Research* **48**, D696–D703.
- COURTIER-ORGOGOZO, V. & MARTIN, A. (2020). The coding loci of evolution and domestication: current knowledge and implications for bio-inspired genome editing. *Journal of Experimental Biology* **223**, jeb208934.
- *CRAWFORD, N. G., KELLY, D. E., HANSEN, M. E. B., BELTRAME, M. H., FAN, S., BOWMAN, S. L., JEWETT, E., RANCIARO, A., THOMPSON, S., LO, Y., PFEIFER, S. P., JENSEN, J. D., CAMPBELL, M. C., BEGGS, W., HORMOZDIARI, F., ET AL. (2017). Loci associated with skin pigmentation identified in African populations. *Science* **358**, eaan8433.
- 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., MAPPES, J., MARSHALL, J., MERRILL, R., OSORIO, D., PRUM, R., ET AL. (2017). The biology of color. *Science* **357**, eaan0221.
- *DAMÉ, M. C. F., XAVIER, G. M., OLIVEIRA-FILHO, J. P., BORGES, A. S., OLIVEIRA, H. N., RIET-CORREA, F. & SCHILD, A. L. (2012). A nonsense mutation in the tyrosinase gene causes albinism in water buffalo. *BMC Genetics* **13**, 62.
- *DAVID, V. A., MENOTTI-RAYMOND, M., WALLACE, A. C., ROELKE, M., KEHLER, J., LEIGHTY, R., EIZIRIK, E., HANNAH, S. S., NELSON, G., SCHÄFFER, A. A., CONNELLY, C. J., O'BRIEN, S. J. & RYUGO, D. K. (2014). Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. *G3* **4**, 1881–1891.
- DEL MARMOL, V. & BEERMANN, F. (1996). Tyrosinase and related proteins in mammalian pigmentation. *FEBS Letters* **381**, 165–168.
- *DOMYAN, E. T., GUERNSEY, M. W., KRONENBERG, Z., KRISHNAN, S., BOISSY, R. E., VICKREY, A. L., RODGERS, C., CASSIDY, P., LEACHMAN, S. A., FONDON, J. W., YANDELL, M. & SHAPIRO, M. D. (2014). Epistatic and combinatorial effects of pigmentary gene mutations in the domestic pigeon. *Current Biology* **24**, 459–464.
- *DORSHORST, B., HENEGAR, C., LIAO, X., SÄLLMAN ALMÉN, M., RUBIN, C.-J., ITO, S., WAKAMATSU, K., STOTHARD, P., VAN DOORMAAL, B., PLASTOW, G., BARSH, G. S. & ANDERSSON, L. (2015). Dominant red coat color in Holstein Cattle is associated with a missense mutation in the coatmer protein complex, subunit alpha (COPA) gene. *PLoS One* **10**, e0128969.
- *DORSHORST, B., MOLIN, A.-M., RUBIN, C.-J., JOHANSSON, A. M., STRÖMSTEDT, L., PHAM, M.-H., CHEN, C.-F., HALLBÖÖK, F., ASHWELL, C. & ANDERSSON, L. (2011). A complex genomic rearrangement involving the endothelin 3 locus causes dermal hyperpigmentation in the chicken. *PLoS Genetics* **7**, e1002412.
- *DOUCET, S. M., SHAWKEY, M. D., RATHBURN, M. K., MAYS, H. L. & MONTGOMERIE, R. (2004). Concordant evolution of plumage colour, feather microstructure and a melanocortin receptor gene between mainland and Island populations of a fairy-wren. *Proceedings. Biological Sciences* **271**, 1663–1670.
- *DREGER, D. L., PARKER, H. G., OSTRANDER, E. A. & SCHMUTZ, S. M. (2013). Identification of a mutation that is associated with the saddle tan and black-and-tan phenotypes in basset hounds and Pembroke welsh corgis. *The Journal of Heredity* **104**, 399–406.
- *DREGER, D. L. & SCHMUTZ, S. M. (2010). A new mutation in MC1R explains a coat color phenotype in 2 “old” breeds: saluki and afghan hound. *Journal of Heredity* **101**, 644–649.
- *DREGER, D. L. & SCHMUTZ, S. M. (2011). A SINE insertion causes the black-and-tan and saddle tan phenotypes in domestic dogs. *The Journal of Heredity* **102**(Suppl 1), S11–S18.
- *DRÖGEMÜLLER, C., GIESE, A., MARTINS-WESS, F., WIEDEMANN, S., ANDERSSON, L., BREINIG, B., FRIES, R. & LEEB, T. (2006). The mutation causing the black-and-tan pigmentation phenotype of Mangalitzia pigs maps to the porcine ASIP locus but does not affect its coding sequence. *Mammalian Genome* **17**, 58–66.
- *DRÖGEMÜLLER, C., PHILIPP, U., HAASE, B., GÜNZEL-APEL, A.-R. & LEEB, T. (2007). A noncoding melanophilin gene (MLPH) SNP at the splice donor of exon 1 represents a candidate causal mutation for coat color dilution in dogs. *The Journal of Heredity* **98**, 468–473.
- *DU, J. & FISHER, D. E. (2002). Identification of Aim-1 as the underwhite mouse mutant and its transcriptional regulation by MITF. *The Journal of Biological Chemistry* **277**, 402–406.
- *DÜRIG, N., JUDE, R., HOLL, H., BROOKS, S. A., LAFAYETTE, C., JAGANNATHAN, V. & LEEB, T. (2017a). Whole genome sequencing reveals a novel deletion variant in the KIT gene in horses with white spotted coat colour phenotypes. *Animal Genetics* **48**, 483–485.
- *DÜRIG, N., JUDE, R., JAGANNATHAN, V. & LEEB, T. (2017b). A novel MITF variant in a white American Standardbred foal. *Animal Genetics* **48**, 123–124.
- *DÜRIG, N., LETKO, A., LEPORI, V., HADJI RASOULIHA, S., LOECHEL, R., KEHL, A., HYTÖNEN, M. K., LOHI, H., MAURI, N., DIETRICH, J., WIEDMER, M., DRÖGEMÜLLER, M., JAGANNATHAN, V., SCHMUTZ, S. M. & LEEB, T. (2018). Two MC1R loss-of-function alleles in cream-coloured Australian cattle dogs and white huskies. *Animal Genetics* **49**, 284–290.
- *DURKIN, K., COPPIETERS, W., DRÖGEMÜLLER, C., AHARIZ, N., CAMBISANO, N., DRUET, T., FASQUELLE, C., HAILE, A., HORIN, P., HUANG, L., KAMATANI, Y., KARIM, L., LATHROP, M., MOSER, S., OLDENBROEK, K., ET AL. (2012). Serial translocation by means of circular intermediates underlies colour sidedness in cattle. *Nature* **482**, 81–84.
- *EDWARDS, M., BIGHAM, A., TAN, J., LI, S., GOZDZIK, A., ROSS, K., JIN, L. & PARRA, E. J. (2010). Association of the OCA2 polymorphism His615Arg with melanin content in east Asian populations: further evidence of convergent evolution of skin pigmentation. *PLoS Genetics* **6**, e1000867.
- *EIBERG, H., TROELSEN, J., NIELSEN, M., MIKKELSEN, A., MENGEL-FROM, J., KJÆR, K. W. & HANSEN, L. (2008). Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. *Human Genetics* **123**, 177–187.
- *EIZIRIK, E., YUHKI, N., JOHNSON, W. E., MENOTTI-RAYMOND, M., HANNAH, S. S. & O'BRIEN, S. J. (2003). Molecular genetics and evolution of melanism in the cat family. *Current Biology* **13**, 448–453.
- *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. & ANDERSSON, L.

- (2021). A multispecies BCO2 beak color polymorphism in the Darwin's finch radiation. *Current Biology* **31**, 5597–5604.e7.
- *FAN, R., GU, Z., GUANG, X., MARÍN, J. C., VARAS, V., GONZÁLEZ, B. A., WHEELER, J. C., HU, Y., LI, E., SUN, X., YANG, X., ZHANG, C., GAO, W., HE, J., MUNCH, K., *ET AL.* (2020). Genomic analysis of the domestication and post-Spanish conquest evolution of the llama and alpaca. *Genome Biology* **21**, 159.
- *FEINER, N., BRUN-USAN, M., ANDRADE, P., PRANTER, R., PARK, S., MENKE, D. B., GENEVA, A. J. & ULLER, T. (2022). A single locus regulates a female-limited color pattern polymorphism in a reptile. *Science Advances* **8**, eabm2387.
- FENG, Y., MCQUILLAN, M. A. & TISHKOFF, S. A. (2021). Evolutionary genetics of skin pigmentation in African populations. *Human Molecular Genetics* **30**, R88–R97.
- *FONTANESI, L., BERETTI, F., RIGGIO, V., DALL'OLIO, S., GONZÁLEZ, E. G., FINOCCHIARO, R., DAVOLI, R., RUSSO, V. & PORTOLANO, B. (2009). Missense and nonsense mutations in melanocortin 1 receptor (MC1R). Gene of different goat breeds: association with red and black coat colour phenotypes but with unexpected evidences. *BMC Genetics* **10**, 47.
- *FONTANESI, L., FORESTIER, L., ALLAIN, D., SCOTTI, E., BERETTI, F., DERETZ-PICOULET, S., PECCHIOLI, E., VERNESI, C., ROBINSON, T. J., MALANEY, J. L., RUSSO, V. & OULMOUDEN, A. (2010a). Characterization of the rabbit agouti signaling protein (ASIP) gene: transcripts and phylogenetic analyses and identification of the causative mutation of the nonagouti black coat colour. *Genomics* **95**, 166–175.
- *FONTANESI, L., SCOTTI, E., ALLAIN, D. & DALL'OLIO, S. (2014). A frameshift mutation in the melanophilin gene causes the dilute coat colour in rabbit (*Oryctolagus cuniculus*) breeds. *Animal Genetics* **45**, 248–255.
- *FONTANESI, L., SCOTTI, E., COLOMBO, M., BERETTI, F., FORESTIER, L., DALL'OLIO, S., DERETZ, S., RUSSO, V., ALLAIN, D. & OULMOUDEN, A. (2010b). A composite six bp in-frame deletion in the melanocortin 1 receptor (MC1R) gene is associated with the Japanese brindling coat colour in rabbits (*Oryctolagus cuniculus*). *BMC Genetics* **11**, 59.
- *FONTANESI, L., SCOTTI, E. & RUSSO, V. (2012). Haplotype variability in the bovine MITF gene and association with piebaldism in Holstein and Simmental cattle breeds. *Animal Genetics* **43**, 250–256.
- *FONTANESI, L., TAZZOLI, M., BERETTI, F. & RUSSO, V. (2006). Mutations in the melanocortin 1 receptor (MC1R) gene are associated with coat colours in the domestic rabbit (*Oryctolagus cuniculus*). *Animal Genetics* **37**, 489–493.
- *FRUDAKIS, T., THOMAS, M., GASKIN, Z., VENKATESWARLU, K., CHANDRA, K. S., GINJUPALLI, S., GUNTURI, S., NATRAJAN, S., PONNUSWAMY, V. K. & PONNUSWAMY, K. N. (2003). Sequences associated with human iris pigmentation. *Genetics* **165**, 2071–2083.
- *FUKAMACHI, S., KINOSHITA, M., TSUJIMURA, T., SHIMADA, A., ODA, S., SHIMA, A., MEYER, A., KAWAMURA, S. & MITANI, H. (2008). Rescue from oculocutaneous albinism type 4 using medaka *slc45a2* cDNA driven by its own promoter. *Genetics* **178**, 761–769.
- *FUKAMACHI, S., SHIMADA, A. & SHIMA, A. (2001). Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka. *Nature Genetics* **28**, 381–385.
- *GANGOSO, L., GRANDE, J. M., DUCREST, A.-L., FIGUEROA, J., BORTOLOTTI, G. R., ANDRÉS, J. A. & ROULIN, A. (2011). MC1R-dependent, melanin-based colour polymorphism is associated with cell-mediated response in the Eleonora's falcon. *Journal of Evolutionary Biology* **24**, 2055–2063.
- GARCÍA-BORRÓN, J. C., SÁNCHEZ-LAORDEN, B. L. & JIMÉNEZ-CERVANTES, C. (2005). Melanocortin-1 receptor structure and functional regulation. *Pigment Cell Research* **18**, 393–410.
- *GAZDA, M. A., TOOMEY, M. B., ARAÚJO, P. M., LOPES, R. J., AFONSO, S., MYERS, C. A., SERRES, K., KISER, P. D., HILL, G. E., CORBO, J. C. & CARNEIRO, M. (2020). Genetic basis of de novo appearance of carotenoid ornamentation in bare parts of canaries. *Molecular Biology and Evolution* **37**, 1317–1328.
- GIBERT, J.-M. (2017). The flexible stem hypothesis: evidence from genetic data. *Development Genes and Evolution* **227**, 297–307.
- *GIRARDOT, M., GUIBERT, S., LAFORET, M.-P., GALLARD, Y., LARROQUE, H. & OULMOUDEN, A. (2006). The insertion of a full-length *Bos taurus* LINE element is responsible for a transcriptional deregulation of the Normande agouti gene. *Pigment Cell Research* **19**, 346–355.
- *GISKA, I., FARELO, L., PIMENTA, J., SEIXAS, F. A., FERREIRA, M. S., MARQUES, J. P., MIRANDA, I., LETTY, J., JENNY, H., HACKLÄNDER, K., MAGNUSSEN, E. & MELO-FERREIRA, J. (2019). Introgression drives repeated evolution of winter coat color polymorphism in hares. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 24150–24156.
- GODA, M. & FUJII, R. (1995). Blue chromatophores in two species of callionymid fish. *Zoological Science* **12**, 811–813.
- *GRAF, J., HODGSON, R. & VAN DAAL, A. (2005). Single nucleotide polymorphisms in the MATP gene are associated with normal human pigmentation variation. *Human Mutation* **25**, 278–284.
- *GRATTEN, J., BERARDI, D., LOWDER, B. V., MCRAE, A. F., VISSCHER, P. M., PEMBERTON, J. M. & SLATE, J. (2007). Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings. Biological Sciences* **274**, 619–626.
- *GRATTEN, J., PILKINGTON, J. G., BROWN, E. A., BERARDI, D., PEMBERTON, J. M. & SLATE, J. (2010). The genetic basis of recessive self-colour pattern in a wild sheep population. *Heredity* **104**, 206–214.
- *GROSS, J. B., BOROWSKY, R. & TABIN, C. J. (2009). A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. *PLoS Genetics* **5**, e1000326.
- *GUENTHER, C. A., TASIC, B., LUO, L., BEDELL, M. A. & KINGSLEY, D. M. (2014). A molecular basis for classic blond hair color in Europeans. *Nature Genetics* **46**, 748–752.
- *GUERNSEY, M. W., RITSCHER, L., MILLER, M. A., SMITH, D. A., SCHÖNEBERG, T. & SHAPIRO, M. D. (2013). A Val85Met mutation in melanocortin-1 receptor is associated with reductions in eumelanin pigmentation and cell surface expression in domestic rock pigeons (*Columba livia*). *PLoS One* **8**, e74475.
- *GUNNARSSON, U., HELLSTRÖM, A. R., TIXIER-BOICHARD, M., MINVIELLE, F., BED'HOM, B., ITO, S., JENSEN, P., RATTINK, A., VEREIJKEN, A. & ANDERSSON, L. (2007). Mutations in SLC45A2 cause plumage color variation in chicken and Japanese quail. *Genetics* **175**, 867–877.
- *GUNNARSSON, U., KERJE, S., BED'HOM, B., SAHLQVIST, A.-S., EKWALL, O., TIXIER-BOICHARD, M., KÄMPE, O. & ANDERSSON, L. (2011). The dark brown plumage color in chickens is caused by an 8.3-kb deletion upstream of SOX10. *Pigment Cell & Melanoma Research* **24**, 268–274.
- *GUO, J., SUN, X., MAO, A., LIU, H., ZHAN, S., LI, L., ZHONG, T., WANG, L., CAO, J., LIU, G. E. & ZHANG, H. (2022). A 13.42-kb tandem duplication at the ASIP locus is strongly associated with the depigmentation phenotype of non-classic Swiss markings in goats. *BMC Genomics* **23**, 437.
- *GUO, L., BLOOM, J., SYKES, S., HUANG, E., KASHIF, Z., PHAM, E., HO, K., ALCARAZ, A., XIAO, X. G., DUARTE-VOGEL, S. & KRUGLYAK, L. (2021). Genetics of white color and iridophoroma in 'lemon frost' leopard geckos. *PLoS Genetics* **17**, e1009580.
- *GUSTAFSON, N. A., GANDOLFI, B. & LYONS, L. A. (2017). Not another type of potato: MC1R and the russet coloration of Burmese cats. *Animal Genetics* **48**, 116–120.
- *GUTIÉRREZ-GIL, B., WIENER, P. & WILLIAMS, J. L. (2007). Genetic effects on coat colour in cattle: dilution of eumelanin and pheomelanin pigments in an F2-backcross Charolais x Holstein population. *BMC Genetics* **8**, 56.
- *HAASE, B., BROOKS, S. A., SCHLUMBAUM, A., AZOR, P. J., BAILEY, E., ALAEDDINE, F., MÉVISSIN, M., BURGER, D., PONCET, P.-A., RIEDER, S. & LEEB, T. (2007). Allelic heterogeneity at the equine KIT locus in dominant white (W) horses. *PLoS Genetics* **3**, e195.
- *HAASE, B., BROOKS, S. A., TOZAKI, T., BURGER, D., PONCET, P.-A., RIEDER, S., HASEGAWA, T., PENEDO, C. & LEEB, T. (2009). Seven novel KIT mutations in horses with white coat colour phenotypes. *Animal Genetics* **40**, 623–629.
- *HAASE, B., JAGANNATHAN, V., RIEDER, S. & LEEB, T. (2015a). A novel KIT variant in an Icelandic horse with white-spotted coat colour. *Animal Genetics* **46**, 466.
- *HAASE, B., RIEDER, S. & LEEB, T. (2015b). Two variants in the KIT gene as candidate causative mutations for a dominant white and a white spotting phenotype in the donkey. *Animal Genetics* **46**, 321–324.
- *HAASE, B., RIEDER, S., TOZAKI, T., HASEGAWA, T., PENEDO, M. C. T., JUDE, R. & LEEB, T. (2011). Five novel KIT mutations in horses with white coat colour phenotypes. *Animal Genetics* **42**, 337–339.
- *HAASE, B., SIGNER-HASLER, H., BINNS, M. M., OBEXER-RUFF, G., HAUSWIRTH, R., BELLONE, R. R., BURGER, D., RIEDER, S., WADE, C. M. & LEEB, T. (2013). Accumulating mutations in series of haplotypes at the KIT and MITF loci are major determinants of white markings in Franches-Montagnes horses. *PLoS One* **8**, e75071.
- HANLY, J. J., LIVRAGHI, L., HERYANTO, C., McMILLAN, W. O., JIGGINS, C. D., GILBERT, L. E. & MARTIN, A. (2021). A macromutation eliminates colour patterning in captive butterflies. *bioRxiv*. <https://doi.org/10.1101/2021.10.29.466422>.
- HARRIS, M. L., BAXTER, L. L., LOFTUS, S. K. & PAVAN, W. J. (2010). Sox proteins in melanocyte development and melanoma. *Pigment Cell & Melanoma Research* **23**, 496–513.
- *HAUSWIRTH, R., HAASE, B., BLATTER, M., BROOKS, S. A., BURGER, D., DRÖGEMÜLLER, C., GERBER, V., HENKE, D., JANDA, J., JUDE, R., MAGDESIAN, K. G., MATTHEWS, J. M., PONCET, P.-A., SVANSSON, V., TOZAKI, T., *ET AL.* (2012). Mutations in MITF and PAX3 cause 'splashed white' and other white spotting phenotypes in horses. *PLoS Genetics* **8**, e1002653.
- *HAUSWIRTH, R., JUDE, R., HAASE, B., BELLONE, R. R., ARCHER, S., HOLL, H., BROOKS, S. A., TOZAKI, T., PENEDO, M. C. T., RIEDER, S. & LEEB, T. (2013). Novel variants in the KIT and PAX3 genes in horses with white-spotted coat colour phenotypes. *Animal Genetics* **44**, 763–765.

- *HÉDAN, B., CADIEU, E., BOTHEREL, N., DUFAYRE DE CITRES, C., LETKO, A., RIMBAULT, M., DRÖGEMÜLLER, C., JAGANNATHAN, V., DERRIEN, T., SCHMUTZ, S., LEEB, T. & ANDRÉ, C. (2019). Identification of a missense variant in MFSD12 involved in dilution of pheomelanin leading to white or cream coat color in dogs. *Genes* **10**, 386.
- HEDRICK, P. (2015). Heterozygote advantage: the effect of artificial selection in livestock and pets. *The Journal of Heredity* **106**, 141–154.
- *HELLSTRÖM, A. R., SUNDSTRÖM, E., GUNNARSSON, U., BED'HOM, B., TIXIER-BOICHARD, M., HONAKER, C. F., SAHLQVIST, A.-S., JENSEN, P., KÄMPE, O., SIEGEL, P. B., KERJE, S. & ANDERSSON, L. (2010). Sex-linked barring in chickens is controlled by the CDKN2A/B tumour suppressor locus. *Pigment Cell & Melanoma Research* **23**, 521–530.
- *HENKEL, J., DUBACHER, A., BANGERTER, E., HERREN, U., AMMANN, P., DRÖGEMÜLLER, C., FLURY, C. & LEEB, T. (2021). Introgression of ASIP and TYRP1 alleles explains coat color variation in Valais goats. *Journal of Heredity* **112**, 452–457.
- *HENKEL, J., LAFAYETTE, C., BROOKS, S. A., MARTIN, K., PATTERSON-ROSA, L., COOK, D., JAGANNATHAN, V. & LEEB, T. (2019). Whole-genome sequencing reveals a large deletion in the MITF gene in horses with white spotted coat colour and increased risk of deafness. *Animal Genetics* **50**, 172–174.
- *HIRAGAKI, T., INOUE-MURAYAMA, M., MIWA, M., FUJIWARA, A., MIZUTANI, M., MINVIELLE, F. & ITO, S. (2008). Recessive black is allelic to the yellow plumage locus in Japanese quail and associated with a frameshift deletion in the ASIP gene. *Genetics* **178**, 771–775.
- *HOBAN, R., CASTLE, K., HAMILTON, N. & HAASE, B. (2018). Novel KIT variants for dominant white in the Australian horse population. *Animal Genetics* **49**, 99–100.
- HOEKSTRA, H. E. (2006). Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity* **97**, 222–234.
- HOEKSTRA, H. E. & COYNE, J. A. (2007). The locus of evolution: Evo Devo and the genetics of adaptation. *Evolution* **61**, 995–1016.
- *HOEKSTRA, H. E., HIRSCHMANN, R. J., BUNDEY, R. A., INSEL, P. A. & CROSSLAND, J. P. (2006). A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**, 101–104.
- *HOFFMAN, J. L., BAUER, E., PAIJMANS, A. J., HUMBLE, E., BECKMANN, L. M., KUBETSCHKE, C., CHRISTALLER, F., KRÖCKER, N., FUCHS, B., MORERAS, A., SHIHLOMULE, Y. D., BESTER, M. N., CLEARY, A. C., DE BRUYN, P. J. N., FORCADA, J., ET AL. (2018). A global cline in a colour polymorphism suggests a limited contribution of gene flow towards the recovery of a heavily exploited marine mammal. *Royal Society Open Science* **5**, 181227.
- *HOLL, H., ISAZA, R., MOHAMOUD, Y., AHMED, A., ALMATHEN, F., YOUSEF, C., GAOUAR, S., ANTZAK, D. F. & BROOKS, S. (2017a). A frameshift mutation in KIT is associated with white spotting in the Arabian Camel. *Genes* **8**, 102.
- *HOLL, H. M., BROOKS, S. A., CARPENTER, M. L., BUSTAMANTE, C. D. & LAFAYETTE, C. (2017b). A novel splice mutation within equine KIT and the W15 allele in the homozygous state lead to all white coat color phenotypes. *Animal Genetics* **48**, 497–498.
- *HOLL, H. M., PFLUG, K. M., YATES, K. M., HOEFS-MARTIN, K., SHEPARD, C., COOK, D. G., LAFAYETTE, C. & BROOKS, S. A. (2019). A candidate gene approach identifies variants in SLC45A2 that explain dilute phenotypes, pearl and sunshine, in compound heterozygote horses. *Animal Genetics* **50**, 271–274.
- HOU, L., PANTHIER, J. J. & ARNHEITER, H. (2000). Signaling and transcriptional regulation in the neural crest-derived melanocyte lineage: interactions between KIT and MITF. *Development* **127**, 5379–5389.
- *HRČKOVA, TURNOVA, E., MAJCHRAKOVÁ, Z., BIELIKOVÁ, M., SOLTYS, K., TURNA, J. & DUDAS, A. (2017). A novel mutation in the TYRP1 gene associated with brown coat colour in the Australian Shepherd Dog Breed. *Animal Genetics* **48**, 626.
- HU, Z.-L., PARK, C. A. & REECY, J. M. (2021). Bringing the Animal QTLdb and CoTDB into the future: meeting new challenges and providing updated services. *Nucleic Acids Research* **50**, D956–D961.
- HUANG, D., LEWIS, V. M., FOSTER, T. N., TOOMEY, M. B., CORBO, J. C. & PARICHY, D. M. (2021). Development and genetics of red coloration in the zebrafish relative *Danio albolineatus*. *eLife* **10**, e70253.
- HUBBARD, J. K., UY, J. A. C., HAUBER, M. E., HOEKSTRA, H. E. & SAFRAN, R. J. (2010). Vertebrate pigmentation: from underlying genes to adaptive function. *Trends in Genetics* **26**, 231–239.
- *HUG, P., JUDE, R., HENKEL, J., JAGANNATHAN, V. & LEEB, T. (2019). A novel KIT deletion variant in a German Riding Pony with white-spotting coat colour phenotype. *Animal Genetics* **50**, 761–763.
- *IMES, D. L., GEARY, L. A., GRAHN, R. A. & LYONS, L. A. (2006). Albinism in the domestic cat (*Felis catus*) is associated with a tyrosinase (TYR) mutation. *Animal Genetics* **37**, 175–178.
- *IMSLAND, F., MCGOWAN, K., RUBIN, C.-J., HENEGAR, C., SUNDSTRÖM, E., BERGLUND, J., SCHWOCHOW, D., GUSTAFSON, U., IMSLAND, P., LINDBLAD-TOH, K., LINDGREN, G., MIKKO, S., MILLON, L., WADE, C., SCHUBERT, M., ET AL. (2016). Regulatory mutations in TBX3 disrupt asymmetric hair pigmentation that underlies Dun camouflage color in horses. *Nature Genetics* **48**, 152–158.
- *ISHIDA, Y., DAVID, V. A., EIZIRIK, E., SCHÄFFER, A. A., NEELAM, B. A., ROELKE, M. E., HANNAH, S. S., O'BRIEN, S. J. & MENOTTI-RAYMOND, M. (2006). A homozygous single-base deletion in MLPH causes the dilute coat color phenotype in the domestic cat. *Genomics* **88**, 698–705.
- *JANSSEN, K. & MUNDY, N. I. (2013). Molecular population genetics of the melanin plumage polymorphism in Arctic skuas (*Stercorarius parasiticus*): evidence for divergent selection on plumage colour. *Molecular Ecology* **22**, 4634–4643.
- *JOHNSON, J. A., AMBERS, A. D. & BURNHAM, K. K. (2012). Genetics of plumage color in the Gyrfalcon (*Falco rusticolus*): analysis of the melanocortin-1 receptor gene. *The Journal of Heredity* **103**, 315–321.
- *JOHNSON, J. L., KOZYSYA, A., KHARLAMOVA, A. V., GULEVICH, R. G., PERELMAN, P. L., FONG, H. W. F., VLADIMIROVA, A. V., OSKINA, I. N., TRUT, L. N. & KUKKOVA, A. V. (2015). Platinum coat color in red fox (*Vulpes vulpes*) is caused by a mutation in an autosomal copy of KIT. *Animal Genetics* **46**, 190–199.
- JOHNSTON, S. E., GRATTEN, J., BERENOS, C., PILKINGTON, J. G., CLUTTON-BROCK, T. H., PEMBERTON, J. M. & SLATE, J. (2013). Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* **502**, 93–95.
- *JOLLY, R. D., WILLS, J. L., KENNY, J. E., CAHILL, J. I. & HOWE, L. (2008). Coat-color dilution and hypochrichosis in Hereford crossbred calves. *New Zealand Veterinary Journal* **56**, 74–77.
- *JONES, M. R., MILLS, L. S., ALVES, P. C., CALLAHAN, C. M., ALVES, J. M., LAFFERTY, D. J. R., JIGGINS, F. M., JENSEN, J. D., MELO-FERREIRA, J. & GOOD, J. M. (2018). Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. *Science* **360**, 1355–1358.
- *KAELIN, C. B., XU, X., HONG, L. Z., DAVID, V. A., MCGOWAN, K. A., SCHMIDT-KÜNTZEL, A., ROELKE, M. E., PINO, J., PONTIUS, J., COOPER, G. M., MANUEL, H., SWANSON, W. F., MARKER, L., HARPER, C. K., VAN DYK, A., ET AL. (2012). Specifying and sustaining pigmentation patterns in domestic and wild cats. *Science* **337**, 1536–1541.
- *KAMBE, Y., TANIKAWA, T., MATSUMOTO, Y., TOMOZAWA, M., APLIN, K. P. & SUZUKI, H. (2011). Origin of agouti-melanistic polymorphism in wild black rats (*Rattus rattus*) inferred from Mc1r gene sequences. *Zoological Science* **28**, 560–567.
- *KARLSSON, E. K., BARANOWSKA, I., WADE, C. M., SALMON HILLBERTZ, N. H. C., ZODY, M. C., ANDERSON, N., BIAGI, T. M., PATTERSON, N., PIELBERG, G. R., KULBOKAS, E. J., COMSTOCK, K. E., KELLER, E. T., MESIROV, J. P., VON EULER, H., KÄMPE, O., ET AL. (2007). Efficient mapping of mendelian traits in dogs through genome-wide association. *Nature Genetics* **39**, 1321–1328.
- KELSH, R. N., HARRIS, M. L., COLANESI, S. & ERICKSON, C. A. (2009). Stripes and belly-spots – a review of pigment cell morphogenesis in vertebrates. *Seminars in Cell & Developmental Biology* **20**, 90–104.
- *KENNY, E. E., TIMPSON, N. J., SIKORA, M., YEE, M.-C., MORENO-ESTRADA, A., ENG, C., HUNTSMAN, S., BURCHARD, E. G., STONEKING, M., BUSTAMANTE, C. D. & MYLES, S. (2012). Melanesian blond hair is caused by an amino acid change in TYRP1. *Science* **336**, 554.
- *KERJE, S., LIND, J., SCHÜTZ, K., JENSEN, P. & ANDERSSON, L. (2003). Melanocortin 1-receptor (MC1R) mutations are associated with plumage colour in chicken. *Animal Genetics* **34**, 241–248.
- *KERJE, S., SHARMA, P., GUNNARSSON, U., KIM, H., BAGCHI, S., FREDRIKSSON, R., SCHÜTZ, K., JENSEN, P., VON HEIJNE, G., OKIMOTO, R. & ANDERSSON, L. (2004). The dominant white, dun and smoky color variants in chicken are associated with insertion/deletion polymorphisms in the PMEL17 gene. *Genetics* **168**, 1507–1518.
- *KERN, J. A., NEWTON, J., BERRYERE, T. G., RUBIN, E. M., CHENG, J.-F., SCHMUTZ, S. M. & BARSH, G. S. (2004). Characterization of the dog Agouti gene and a nonagouti mutation in German Shepherd Dogs. *Mammalian Genome* **15**, 798–808.
- *KIJAS, J. M., MOLLER, M., PLASTOW, G. & ANDERSSON, L. (2001). A frameshift mutation in MC1R and a high frequency of somatic reversions cause black spotting in pigs. *Genetics* **158**, 779–785.
- *KIJAS, J. M., WALES, R., TÖRNSTEN, A., CHARDON, P., MOLLER, M. & ANDERSSON, L. (1998). Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics* **150**, 1177–1185.
- *KINGSLEY, E. P., MANCEAU, M., WILEY, C. D. & HOEKSTRA, H. E. (2009). Melanism in *Peromyscus* is caused by independent mutations in agouti. *PLoS One* **4**, e6435.
- *KINOSHITA, K., AKIYAMA, T., MIZUTANI, M., SHINOMIYA, A., ISHIKAWA, A., YOUNIS, H. H., TSUZUKI, M., NAMIKAWA, T. & MATSUDA, Y. (2014). Endothelin receptor B2 (EDNRB2) is responsible for the tyrosinase-independent recessive white (mo(w)) and mottled (mo) plumage phenotypes in the chicken. *PLoS One* **9**, e86361.
- *KLUNGLAND, H., VÅGE, D., GOMEZ-RAYA, L., ADALSTEINSSON, S. & LIEN, S. (1995). The role of melanocyte-stimulated hormone (MSH) receptor in bovine coat color determination. *Mammalian Genome* **6**, 636–639.

- *KNIEF, U., BOSSU, C. M., SAINO, N., HANSSON, B., POELSTRA, J., VIJAY, N., WEISSENSTEINER, M. & WOLF, J. B. W. (2019). Epistatic mutations under divergent selection govern phenotypic variation in the crow hybrid zone. *Nature Ecology & Evolution* **3**, 570–576.
- *KOGA, A. & HORI, H. (1997). Albinism due to transposable element insertion in fish. *Pigment Cell Research* **10**, 377–381.
- *KON, T., OMORI, Y., FUKUTA, K., WADA, H., WATANABE, M., CHEN, Z., IWASAKI, M., MISHINA, T., MATSUZAKI, S.-I. S., YOSHIHARA, D., ARAKAWA, J., KAWAKAMI, K., TOYODA, A., BURGESS, S. M., NOGUCHI, H., *ET AL.* (2020). The genetic basis of morphological diversity in domesticated goldfish. *Current Biology* **30**, 2260–2274.e6.
- *KRATOCHWIL, C. F., KAUTT, A. F., NATER, A., HÄRER, A., LIANG, Y., HENNING, F. & MEYER, A. (2022). An intronic transposon insertion associates with a trans-species color polymorphism in Midas cichlid fishes. *Nature Communications* **13**, 296.
- *KRATOCHWIL, C. F., LIANG, Y., GERWIN, J., WOLTERING, J. M., URBAN, S., HENNING, F., MACHADO-SCHIAFFINO, G., HULSEY, C. D. & MEYER, A. (2018). Agouti-related peptide 2 facilitates convergent evolution of stripe patterns across cichlid fish radiations. *Science* **362**, 457–460.
- KRATOCHWIL, C. F. & MEYER, A. (2015). Closing the genotype–phenotype gap: emerging technologies for evolutionary genetics in ecological model vertebrate systems. *BioEssays* **37**, 213–226.
- *KRATOCHWIL, C. F., URBAN, S. & MEYER, A. (2019). Genome of the Malawi golden cichlid fish (*Melanochromis auratus*) reveals loss of *oca2* in an amelanistic morph. *Pigment Cell & Melanoma Research* **32**, 719–723.
- KRONFORST, M. R., BARSH, G. S., KOPP, A., MALLET, J., MONTEIRO, A., MULLEN, S. P., PROTAS, M., ROSENBLUM, E. B., SCHNEIDER, C. J. & HOEKSTRA, H. E. (2012). Unraveling the thread of nature's tapestry: the genetics of diversity and convergence in animal pigmentation. *Pigment Cell & Melanoma Research* **25**, 411–433.
- *KUKKOVA, A. V., JOHNSON, J. L., KHARLAMOVA, A. V., VLADIMIROVA, A. V., SHEPELEVA, D. V., GULEVICH, R. G. & TRUT, L. N. (2016). Georgian white coat color of red fox (*Vulpes vulpes*) maps to fox chromosome 2 in the region containing KIT gene. *Animal Genetics* **47**, 514–515.
- *KÜTTTEL, L., LETKO, A., HÄFLIGER, I. M., SIGNER-HASLER, H., JOLLER, S., HIRSBRUNNER, G., MÉSZÁROS, G., SÖLKNER, J., FLURY, C., LEEB, T. & DRÖGEMÜLLER, C. (2019). A complex structural variant at the KIT locus in cattle with the Pinzgauer spotting pattern. *Animal Genetics* **50**, 423–429.
- *KWON, B. S., HALABAN, R., PONNAZHAGAN, S., KIM, K., CHINTAMANANI, C., BENNETT, D. & PICKARD, R. T. (1995). Mouse silver mutation is caused by a single base insertion in the putative cytoplasmic domain of Pmel 17. *Nucleic Acids Research* **23**, 154–158.
- *LALUEZA-FOX, C., RÖMPLER, H., CARAMELLI, D., STÄUBERT, C., CATALANO, G., HUGHES, D., ROHLAND, N., PILLI, E., LONGO, L., CONDEMI, S., DE LA RASILLA, M., FORTEA, J., ROSAS, A., STONEKING, M., SCHÖNEBERG, T., *ET AL.* (2007). A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. *Science* **318**, 1453–1455.
- *LAMASON, R. L., MOHIDEEN, M.-A. P. K., MEST, J. R., WONG, A. C., NORTON, H. L., AROS, M. C., JURYNAC, M. J., MAO, X., HUMPHREVILLE, V. R., HUMBERT, J. E., SINHA, S., MOORE, J. L., JAGADEESWARAN, P., ZHAO, W., NING, G., *ET AL.* (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* **310**, 1782–1786.
- *LAMICHHANEY, S., FAN, G., WIDEMO, F., GUNNARSSON, U., THALMANN, D. S., HOEPPNER, M. P., KERJE, S., GUSTAFSON, U., SHI, C., ZHANG, H., CHEN, W., LIANG, X., HUANG, L., WANG, J., LIANG, E., *ET AL.* (2016). Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nature Genetics* **48**, 84–88.
- *LANGEVIN, M., SYNKOVA, H., JANCUSKOVA, T. & PEKOVA, S. (2018). Merle phenotypes in dogs-SILV SINE insertions from Mc to Mh. *PLoS One* **13**, e0198536.
- LAPEDRIZA, A., PETRATOU, K. & KELSH, R. N. (2014). Chapter 14 – Neural crest cells and pigmentation. In *Neural Crest Cells* (ed. P. A. TRAINOR), pp. 287–311. Academic Press, Boston.
- LE, L., ESCOBAR, I. E., HO, T., LEFKOVITH, A. J., LATTERI, E., HALTAUFDERHYDE, K. D., DENNIS, M. K., PLOWRIGHT, L., SVIDERSKAYA, E. V., BENNETT, D. C., OANCEA, E. & MARKS, M. S. (2020). SLC45A2 protein stability and regulation of melanosome pH determine melanocyte pigmentation. *Molecular Biology of the Cell* **31**, 2687–2702.
- *LEHNER, S., GÄHLE, M., DIERKS, C., STELTER, R., GERBER, J., BREHM, R. & DISTL, O. (2013). Two-exon skipping within MLPH is associated with coat color dilution in rabbits. *PLoS One* **8**, e84525.
- LENFER, J., NICHOLAS, F. W., CASTLE, K., RAO, A., GREGORY, S., POIDINGER, M., MAILMAN, M. D. & RANGANATHAN, S. (2006). OMIA (Online Mendelian Inheritance in Animals): an enhanced platform and integration into the Entrez search interface at NCBI. *Nucleic Acids Research* **34**, D599–D601.
- *LETKO, A., AMMANN, B., JAGANNATHAN, V., HENKEL, J., LEUTHARD, F., SCHELLING, C., CARNEIRO, M., DRÖGEMÜLLER, C. & LEEB, T. (2020). A deletion spanning the promoter and first exon of the hair cycle-specific ASIP transcript isoform in black and tan rabbits. *Animal Genetics* **51**, 137–140.
- LEWIS, V. M., SAUNDERS, L. M., LARSON, T. A., BAIN, E. J., STURIALE, S. L., GUR, D., CHOWDHURY, S., FLYNN, J. D., ALLEN, M. C., DEHEYN, D. D., LEE, J. C., SIMON, J. A., LIPPINCOTT-SCHWARTZ, J., RAIBLE, D. W. & PARICHY, D. M. (2019). Fate plasticity and reprogramming in genetically distinct populations of *Danio* leucophores. *Proceedings of the National Academy of Sciences* **116**, 11806–11811.
- *LI, J., BED'HOM, B., MARTHEY, S., VALADE, M., DUREUX, A., MOROLDO, M., PÉCHOUX, C., COVILLE, J.-L., GOURICHON, D., VIEAUD, A., DORSHORST, B., ANDERSSON, L. & TIXIER-BOICHARD, M. (2019a). A missense mutation in TYRPI causes the chocolate plumage color in chicken and alters melanosome structure. *Pigment Cell & Melanoma Research* **32**, 381–390.
- *LI, J., DAVIS, B. W., JERN, P., DORSHORST, B. J., SIEGEL, P. B. & ANDERSSON, L. (2019b). Characterization of the endogenous retrovirus insertion in CYP19A1 associated with henry feathering in chicken. *Mobile DNA* **10**, 38.
- *LI, L., LI, D., LIU, L., LI, S., FENG, Y., PENG, X. & GONG, Y. (2015). Endothelin receptor B2 (EDNRB2) gene is associated with spot plumage pattern in domestic ducks (*Anas platyrhynchos*). *PLoS One* **10**, e0125883.
- *LI, W., SARTELET, A., TAMMA, N., COPPIETERS, W., GEORGES, M. & CHARLIER, C. (2016). Reverse genetic screen for loss-of-function mutations uncovers a frameshifting deletion in the melanophilin gene accountable for a distinctive coat color in Belgian Blue cattle. *Animal Genetics* **47**, 110–113.
- *LI, Y., GENG, X., BAO, L., ELASWAD, A., HUGGINS, K. W., DUNHAM, R. & LIU, Z. (2017). A deletion in the Hermansky-Pudlak syndrome 4 (Hps4) gene appears to be responsible for albinism in channel catfish. *Molecular Genetics and Genomics* **292**, 663–670.
- *LIANG, D., ZHAO, P., SI, J., FANG, L., PAIRO-CASTINEIRA, E., HU, X., XU, Q., HOU, Y., GONG, Y., LIANG, Z., TIAN, B., MAO, H., YINDEE, M., FARUQUE, M. O., KONGVONGXAY, S., *ET AL.* (2021). Genomic analysis revealed a convergent evolution of LINE-1 in coat color: a case study in water buffaloes (*Bubalus bubalis*). *Molecular Biology and Evolution* **38**, 1122–1136.
- *LIM, H. T., ZHONG, T., CHO, I. C., SEO, B. Y., KIM, J. H., LEE, S. S., KO, M. S., PARK, H. B., KIM, B. W., LEE, J. H. & JEON, J. T. (2011). Novel alternative splicing by exon skipping in KIT associated with whole-body roan in an intercrossed population of Landrace and Korean Native pigs. *Animal Genetics* **42**, 451–455.
- *LINNEN, C. R., POH, Y.-P., PETERSON, B. K., BARRETT, R. D. H., LARSON, J. G., JENSEN, J. D. & HOEKSTRA, H. E. (2013). Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339**, 1312–1316.
- *LIU, F., VISSER, M., DUFFY, D. L., HYSI, P. G., JACOBS, L. C., LAO, O., ZHONG, K., WALSH, S., CHAITANYA, L., WOLLSTEIN, A., ZHU, G., MONTGOMERY, G. W., HENDERS, A. K., MANGINO, M., GLASS, D., *ET AL.* (2015). Genetics of skin color variation in Europeans: genome-wide association studies with functional follow-up. *Human Genetics* **134**, 823–835.
- LONA-DURAZO, F., HERNANDEZ-PACHECO, N., FAN, S., ZHANG, T., CHOI, J., KOVACS, M. A., LOFTUS, S. K., LE, P., EDWARDS, M., FORTES-LIMA, C. A., ENG, C., HUNTSMAN, S., HU, D., GÓMEZ-CABEZAS, E. J., MARÍN-PADRÓN, L. C., *ET AL.* (2019). Meta-analysis of GWA studies provides new insights on the genetic architecture of skin pigmentation in recently admixed populations. *BMC Genetics* **20**, 59.
- *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. & CARNEIRO, M. (2016). Genetic basis for red coloration in birds. *Current Biology* **26**, 1427–1434.
- *LYONS, L. A., IMES, D. L., RAH, H. C. & GRAHN, R. A. (2005). Tyrosinase mutations associated with Siamese and Burmese patterns in the domestic cat (*Felis catus*). *Animal Genetics* **36**, 119–126.
- *MACK, M., KOWALSKI, E., GRAHN, R., BRAS, D., PENEDO, M. C. T. & BELLONE, R. (2017). Two variants in SLC24A5 are associated with 'Tiger-Eye' iris pigmentation in Puerto Rican Paso Fino Horses. *G3* **7**, 2799–2806.
- *MACLARY, E. T., PHILLIPS, B., WAUER, R., BOER, E. F., BRUDERS, R., GILVARRY, T., HOLT, C., YANDELL, M. & SHAPIRO, M. D. (2021). Two genomic loci control three eye colors in the domestic pigeon (*Columba livia*). *Molecular Biology and Evolution* **38**, 5376–5390.
- *MAE, Y., NAGARA, K., MIYAZAKI, M., KATSURA, Y., ENOMOTO, Y. & KOGA, A. (2020). Complex intragenic deletion leads to oculocutaneous albinism in tanuki (Japanese raccoon dog). *Genome* **63**, 517–523.
- *MAGDESIAN, K. G., TANAKA, J. & BELLONE, R. R. (2020). A de novo MITF deletion explains a novel splashed white phenotype in an American Paint Horse. *The Journal of Heredity* **111**, 287–293.
- MALLARINO, R., LINDEN, T. A., LINNEN, C. R. & HOEKSTRA, H. E. (2017). The role of isoforms in the evolution of cryptic coloration in *Peromyscus* mice. *Molecular Ecology* **26**, 245–258.
- *MANCEAU, M., DOMINGUES, V. S., MALLARINO, R. & HOEKSTRA, H. E. (2011). The developmental role of agouti in color pattern evolution. *Science* **331**, 1062–1065.

- *MARIAT, D., TAOURIT, S. & GUÉRIN, G. (2003). A mutation in the MATP gene causes the cream coat colour in the horse. *Genetics, Selection, Evolution* **35**, 119–133.
- *MARKLUND, L., MOLLER, M. J., SANDBERG, K. & ANDERSSON, L. (1996). A missense mutation in the gene for melanocyte-stimulating hormone receptor (MC1R) is associated with the chestnut coat color in horses. *Mammalian Genome* **7**, 895–899.
- *MARKLUND, S., KIJAS, J., RODRIGUEZ-MARTINEZ, H., RÖNNSTRAND, L., FUNA, K., MOLLER, M., LANGE, D., EDFORS-LILJA, I. & ANDERSSON, L. (1998). Molecular basis for the dominant white phenotype in the domestic pig. *Genome Research* **8**, 826–833.
- MARTIN, A. & COURTIER-ORGOGOZO, V. (2017). Morphological evolution repeatedly caused by mutations in signaling ligand genes. In *Diversity and Evolution of Butterfly Wing Patterns: An Integrative Approach* (eds T. SEKIMURA and H. F. NIJHOUT), pp. 59–87. Springer, Singapore.
- MARTIN, A. & ORGOGOZO, V. (2013). The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* **67**, 1235–1250.
- MATESIC, L. E., YIP, R., REUSS, A. E., SWING, D. A., O'SULLIVAN, T. N., FLETCHER, C. F., COPELAND, N. G. & JENKINS, N. A. (2001). Mutations in *Mplh*, encoding a member of the Rab effector family, cause the melanosome transport defects observed in *leaden* mice. *Proceedings of the National Academy of Sciences* **98**, 10238–10243.
- MCNAMARA, M. E., ROSSI, V., SLATER, T. S., ROGERS, C. S., DUCREST, A.-L., DUBEY, S. & ROULIN, A. (2021). Decoding the evolution of melanin in vertebrates. *Trends in Ecology & Evolution* **36**, 430–443.
- *MCROBIE, H., THOMAS, A. & KELLY, J. (2009). The genetic basis of melanism in the gray squirrel (*Sciurus carolinensis*). *Journal of Heredity* **100**, 709–714.
- *MCROBIE, H. R., MONCRIEF, N. D. & MUNDT, N. I. (2019). Multiple origins of melanism in two species of North American tree squirrel (*Sciurus*). *BMC Evolutionary Biology* **19**, 140.
- *MENZI, F., KELLER, I., REBER, I., BECK, J., BREINIG, B., SCHÜTZ, E., LEEB, T. & DRÖGEMÜLLER, C. (2016). Genomic amplification of the caprine EDNRA locus might lead to a dose dependent loss of pigmentation. *Scientific Reports* **6**, 28438.
- *METALLINOS, D. L., BOWLING, A. T. & RINE, J. (1998). A missense mutation in the endothelin-B receptor gene is associated with Lethal White Foal Syndrome: an equine version of Hirschsprung disease. *Mammalian Genome* **9**, 426–431.
- *MILLER, C. T., BELEZA, S., POLLEN, A. A., SCHLUTER, D., KITTLES, R. A., SHRIVER, M. D. & KINGSLEY, D. M. (2007). cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* **131**, 1179–1189.
- *MILLER, S. M., GUTHRIE, A. J. & HARPER, C. K. (2016). Single base-pair deletion in ASIP exon 3 associated with recessive black phenotype in impala (*Aepyceros melampus*). *Animal Genetics* **47**, 511–512.
- *MINVIELLE, F., BED'HOM, B., COVILLE, J.-L., ITO, S., INOUE-MURAYAMA, M. & GOURICHON, D. (2010). The 'silver' Japanese quail and the MITF gene: causal mutation, associated traits and homology with the 'blue' chicken plumage. *BMC Genetics* **11**, 15.
- *MINVIELLE, F., CECCHI, T., PASSAMONTI, P., GOURICHON, D. & RENIERI, C. (2009). Plumage colour mutations and melanins in the feathers of the Japanese quail: a first comparison. *Animal Genetics* **40**, 971–974.
- MIURA, I., TAGAMI, M., FUJITANI, T. & OGATA, M. (2018). Spontaneous tyrosinase mutations identified in albinos of three wild frog species. *Genes & Genetic Systems* **92**, 189–196.
- *MIWA, M., INOUE-MURAYAMA, M., AOKI, H., KUNISADA, T., HIRAGAKI, T., MIZUTANI, M. & ITO, S. (2007). Endothelin receptor B2 (EDNRB2) is associated with the panda plumage colour mutation in Japanese quail. *Animal Genetics* **38**, 103–108.
- *MONTAGUE, M. J., LI, G., GANDOLFI, B., KHAN, R., AKEN, B. L., SEARLE, S. M. J., MINX, P., HILLIER, L. W., KOBOLDT, D. C., DAVIS, B. W., DRISCOLL, C. A., BARR, C. S., BLACKSTONE, K., QUILEZ, J., LORENTE-GALDOS, B., ET AL. (2014). Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 17230–17235.
- *MORGAN, H. D., SUTHERLAND, H. G., MARTIN, D. I. & WHITELAW, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nature Genetics* **23**, 314–318.
- MORT, R. L., JACKSON, I. J. & PATTON, E. E. (2015). The melanocyte lineage in development and disease. *Development* **142**, 620–632.
- *MUNDT, N. I., BADCOCK, N. S., HART, T., SCRIBNER, K., JANSSEN, K. & NADEAU, N. J. (2004). Conserved genetic basis of a quantitative plumage trait involved in mate choice. *Science* **303**, 1870–1873.
- *MUNDT, N. I., STAPLEY, J., BENNISON, C., TUCKER, R., TWYMAN, H., KIM, K.-W., BURKE, T., BIRKHEAD, T. R., ANDERSSON, S. & SLATE, J. (2016). Red carotenoid coloration in the zebra finch is controlled by a cytochrome P450 gene cluster. *Current Biology* **26**, 1435–1440.
- *NACHMAN, M. W., HOEKSTRA, H. E. & D'AGOSTINO, S. L. (2003). The genetic basis of adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5268–5273.
- *NADEAU, N. J., MINVIELLE, F., ITO, S., INOUE-MURAYAMA, M., GOURICHON, D., FOLLETT, S. A., BURKE, T. & MUNDT, N. I. (2008). Characterization of Japanese quail yellow as a genomic deletion upstream of the avian homolog of the mammalian ASIP (agouti) gene. *Genetics* **178**, 777–786.
- *NADEAU, N. J., MINVIELLE, F. & MUNDT, N. I. (2006). Association of a Glu92Lys substitution in MC1R with extended brown in Japanese quail (*Coturnix japonica*). *Animal Genetics* **37**, 287–289.
- *NADEAU, N. J., MUNDT, N. I., GOURICHON, D. & MINVIELLE, F. (2007). Association of a single-nucleotide substitution in TYRP1 with roux in Japanese quail (*Coturnix japonica*). *Animal Genetics* **38**, 609–613.
- *NEWTON, J. M., WILKIE, A. L., HE, L., JORDAN, S. A., METALLINOS, D. L., HOLMES, N. G., JACKSON, I. J. & BARSH, G. S. (2000). Melanocortin 1 receptor variation in the domestic dog. *Mammalian Genome* **11**, 24–30.
- *NORRIS, B. J. & WHAN, V. A. (2008). A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Research* **18**, 1282–1293.
- ORTEU, A. & JIGGINS, C. D. (2020). The genomics of coloration provides insights into adaptive evolution. *Nature Reviews Genetics* **21**, 461–475.
- PÁL, C., PAPP, B. & LERCHER, M. J. (2006). An integrated view of protein evolution. *Nature Reviews Genetics* **7**, 337–348.
- *PARIS, J. M., LETKO, A., HÄFLIGER, I. M., AMMANN, P., FLURY, C. & DRÖGEMÜLLER, C. (2019). Identification of two TYRP1 loss-of-function alleles in Valais Red sheep. *Animal Genetics* **50**, 778–782.
- PARICHY, D. M. (2021). Evolution of pigment cells and patterns: recent insights from teleost fishes. *Current Opinion in Genetics & Development* **69**, 88–96.
- PATTERSON, L. B. & PARICHY, D. M. (2013). Interactions with iridophores and the tissue environment required for patterning melanophores and xanthophores during zebrafish adult pigment stripe formation. *PLoS Genetics* **9**, e1003561.
- *PATTERSON, L. B., BAIN, E. J. & PARICHY, D. M. (2014). Pigment cell interactions and differential xanthophore recruitment underlying zebrafish stripe reiteration and *Danio* pattern evolution. *Nature Communications* **5**, 5299.
- PFEIFER, S. P., LAURENT, S., SOUSA, V. C., LINNEN, C. R., FOLL, M., EXCOFFIER, L., HOEKSTRA, H. E. & JENSEN, J. D. (2018). The evolutionary history of Nebraska deer mice: local adaptation in the face of strong gene flow. *Molecular Biology and Evolution* **35**, 792–806.
- *PODOBNIK, M., FROHNHÖFER, H. G., DOOLEY, C. M., ESKOVA, A., NÜSSLEIN-VOLHARD, C. & IRION, U. (2020). Evolution of the potassium channel gene *Kcnj13* underlies colour pattern diversification in *Danio* fish. *Nature Communications* **11**, 6230.
- *POLANOWSKI, A. M., ROBINSON-LAVERICK, S. M., PATON, D. & JARMAN, S. N. (2012). Variation in the tyrosinase gene associated with a white humpback whale (*Megaptera novaeangliae*). *The Journal of Heredity* **103**, 130–133.
- *PRAETORIUS, C., GRILL, C., STACEY, S. N., METCALF, A. M., GORKIN, D. U., ROBINSON, K. C., VAN OTTERLOO, E., KIM, R. S. Q., BERGSTEINSDOTTIR, K., OGMUNDSDOTTIR, M. H., MAGNUSDOTTIR, E., MISHRA, P. J., DAVIS, S. R., GUO, T., ZAIDI, M. R., ET AL. (2013). A polymorphism in IRF4 affects human pigmentation through a tyrosinase-dependent MITF/TFAP2A pathway. *Cell* **155**, 1022–1033.
- *PROTAS, M. E., HERSEY, C., KOCHANEK, D., ZHOU, Y., WILKENS, H., JEFFERY, W. R., ZON, L. I., BOROWSKY, R. & TABIN, C. J. (2006). Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics* **38**, 107–111.
- PROTAS, M. E. & PATEL, N. H. (2008). Evolution of coloration patterns. *Annual Review of Cell and Developmental Biology* **24**, 425–446.
- *QUILLEN, E. E., BAUCHET, M., BIGHAM, A. W., DELGADO-BURBANO, M. E., FAUST, F. X., KLIMENTIDIS, Y. C., MAO, X., STONEKING, M. & SHRIVER, M. D. (2012). OPRM1 and EGFR contribute to skin pigmentation differences between Indigenous Americans and Europeans. *Human Genetics* **131**, 1073–1080.
- *RAADSMAN, H. W., JONAS, E., FLEET, M. R., FULLARD, K., GONGORA, J., CAVANAGH, C. R., TAMMEN, I. & THOMSON, P. C. (2013). QTL and association analysis for skin and fibre pigmentation in sheep provides evidence of a major causative mutation and epistatic effects. *Animal Genetics* **44**, 547–559.
- *REINER, G., TRAMBEREND, K., NIETZELD, F., VOLMER, K., WURMSER, C., FRIES, R. & WILLEMS, H. (2020). A genome-wide scan study identifies a single nucleotide substitution in the tyrosinase gene associated with white coat colour in a red deer (*Cervus elaphus*) population. *BMC Genetics* **21**, 14.
- *REINSON, N., THOMSEN, H., XU, N., BRINK, M., LOOFT, C., KALM, E., BROCKMANN, G. A., GRUPE, S., KÜHN, C., SCHWERR, M., LEYHE, B., HIENDLEDER, S., ERHARDT, G., MEDJUGORAC, I., RUSS, I., ET AL. (1999). A QTL for the degree of spotting in cattle shows synteny with the KIT locus on chromosome 6. *The Journal of Heredity* **90**, 629–634.
- *RIEDER, S., TAOURIT, S., MARIAT, D., LANGLOIS, B. & GUÉRIN, G. (2001). Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (*Equus caballus*). *Mammalian Genome* **12**, 450–455.

- *RITLAND, K., NEWTON, C. & MARSHALL, H. D. (2001). Inheritance and population structure of the white-phased 'Kermode' black bear. *Current Biology* **11**, 1468–1472.
- *ROBBINS, L. S., NADEAU, J. H., JOHNSON, K. R., KELLY, M. A., ROSELLI-REHFUSS, L., BAACK, E., MOUNTJOY, K. G. & CONE, R. D. (1993). Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell* **72**, 827–834.
- *ROBERTS, R. B., MOORE, E. C. & KOCHER, T. D. (2017). An allelic series at *pax7a* is associated with colour polymorphism diversity in Lake Malawi cichlid fish. *Molecular Ecology* **26**, 2625–2639.
- *ROBERTS, R. B., SER, J. R. & KOCHER, T. D. (2009). Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *Science* **326**, 998–1001.
- *ROBIC, A., MORISSON, M., LEROUX, S., GOURICHON, D., VIGNAL, A., THEBAULT, N., FILLON, V., MINVIELLE, F., BED'HOM, B., ZERJAL, T. & PITEF, F. (2019). Two new structural mutations in the 5' region of the ASIP gene cause diluted feather color phenotypes in Japanese quail. *Genetics, Selection, Evolution* **51**, 12.
- ROCKMAN, M. V. (2012). The Qtn program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**, 1–17.
- *ROSENBLUM, E. B., HOEKSTRA, H. E. & NACHMAN, M. W. (2004). Adaptive reptile color variation and the evolution of the *Mclr* gene. *Evolution; International Journal of Organic Evolution* **58**, 1794–1808.
- *ROSENBLUM, E. B., RÖMPLER, H., SCHÖNEBERG, T. & HOEKSTRA, H. E. (2010). Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 2113–2117.
- *ROSENGREN PIELBERG, G., GOLOVKO, A., SUNDRÖM, E., CURIK, I., LENNARTSSON, J., SELTENHAMMER, M. H., DRUML, T., BINNS, M., FITZSIMMONS, C., LINDGREN, G., SANDBERG, K., BAUMUNG, R., VETTERLEIN, M., STRÖMBERG, S., GRABHERR, M., ET AL. (2008). A cis-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse. *Nature Genetics* **40**, 1004–1009.
- *ROTHAMMER, S., KUNZ, E., SEICHTER, D., KREBS, S., WASSERTHEURER, M., FRIES, R., BREM, G. & MEDUGORAC, I. (2017). Detection of two non-synonymous SNPs in *SLC45A2* on *BTA20* as candidate causal mutations for oculocutaneous albinism in Braunvieh cattle. *Genetics, Selection, Evolution* **49**, 73.
- *RUBIN, C.-J., MEGENS, H.-J., MARTINEZ BARRIO, A., MAQBOOL, K., SAYYAB, S., SCHWOCHOW, D., WANG, C., CARLBORG, Ö., JERN, P., JØRGENSEN, C. B., ARCHIBALD, A. L., FREDHOLM, M., GROENEN, M. A. M. & ANDERSSON, L. (2012). Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 19529–19536.
- RUDH, A. & QVARNSTRÖM, A. (2013). Adaptive colouration in amphibians. *Seminars in Cell & Developmental Biology* **24**, 553–561.
- *SAENKO, S. V., LAMICHANEY, S., BARRIO, A. M., RAFATI, N., ANDERSSON, L. & MILINKOVITCH, M. C. (2015). Amelanism in the corn snake is associated with the insertion of an LTR-retrotransposon in the *OCA2* gene. *Scientific Reports* **5**, 17118.
- *SALMON HILLBERTZ, N. H. C., ISAKSSON, M., KARLSSON, E. K., HELLMÉN, E., PIELBERG, G. R., SAVOLAINEN, P., WADE, C. M., VON EULER, H., GUSTAFSON, U., HEDHAMMAR, A., NILSSON, M., LINDBLAD-TOH, K., ANDERSSON, L. & ANDERSSON, G. (2007). Duplication of *FGF3*, *FGF4*, *FGF19* and *ORAOV1* causes hair ridge and predisposition to dermoid sinus in Ridgeback dogs. *Nature Genetics* **39**, 1318–1320.
- *SAN-JOSE, L. M., DUCRET, A.-L., DUCRET, V., BÉZIERS, P., SIMON, C., WAKAMATSU, K. & ROULIN, A. (2015). Effect of the *MC1R* gene on sexual dimorphism in melanin-based colorations. *Molecular Ecology* **24**, 2794–2808.
- SARANATHAN, V. & FINET, C. (2021). Cellular and developmental basis of avian structural coloration. *Current Opinion in Genetics & Development* **69**, 56–64.
- SARKAR, A. & HOCHEDLINGER, K. (2013). The Sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell* **12**, 15–30.
- *SCHMIDT-KÜNTZEL, A., EIZIRIK, E., O'BRIEN, S. J. & MENOTTI-RAYMOND, M. (2005). Tyrosinase and tyrosinase related protein 1 alleles specify domestic cat coat color phenotypes of the albino and brown loci. *The Journal of Heredity* **96**, 289–301.
- *SCHMUTZ, S. M., BERRYERE, T. G., CIOBANU, D. C., MILEHAM, A. J., SCHMIDT, B. H. & FREDHOLM, M. (2004). A form of albinism in cattle is caused by a tyrosinase frameshift mutation. *Mammalian Genome* **15**, 62–67.
- *SCHMUTZ, S. M., BERRYERE, T. G., ELLINWOOD, N. M., KERNS, J. A. & BARSH, G. S. (2003). *MC1R* studies in dogs with melanistic mask or brindle patterns. *Journal of Heredity* **94**, 69–73.
- *SCHMUTZ, S. M., BERRYERE, T. G. & GOLDFINCH, A. D. (2002). *TYRP1* and *MC1R* genotypes and their effects on coat color in dogs. *Mammalian Genome* **13**, 380–387.
- *SCHNEIDER, A., DAVID, V. A., JOHNSON, W. E., O'BRIEN, S. J., BARSH, G. S., MENOTTI-RAYMOND, M. & EIZIRIK, E. (2012). How the leopard hides its spots: ASIP mutations and melanism in wild cats. *PLoS One* **7**, e50386.
- *SCHNEIDER, A., HENEGAR, C., DAY, K., ABSHER, D., NAPOLITANO, C., SILVEIRA, L., DAVID, V. A., O'BRIEN, S. J., MENOTTI-RAYMOND, M., BARSH, G. S. & EIZIRIK, E. (2015). Recurrent evolution of melanism in South American felids. *PLoS Genetics* **11**, e1004892.
- *SCHWOCHOW THALMANN, D., RING, H., SUNDRÖM, E., CAO, X., LARSSON, M., KERJE, S., HÖGLUND, A., FOGELHOLM, J., WRIGHT, D., JEMTH, P., HALLBÖÖK, F., BED'HOM, B., DORSHORST, B., TIXIER-BOIGCHARD, M. & ANDERSSON, L. (2017). The evolution of sex-linked barring alleles in chickens involves both regulatory and coding changes in *CDKN2A*. *PLoS Genetics* **13**, e1006665.
- *SEITZ, J. J., SCHMUTZ, S. M., THUE, T. D. & BUCHANAN, F. C. (1999). A missense mutation in the bovine *MGF* gene is associated with the roan phenotype in Belgian Blue and Shorthorn cattle. *Mammalian Genome* **10**, 710–712.
- *SEVANE, N., SANZ, C. R. & DUNNER, S. (2019). Explicit evidence for a missense mutation in exon 4 of *SLC45A2* gene causing the pearl coat dilution in horses. *Animal Genetics* **50**, 275–278.
- SHEFCHER, K. A., HARRIS, N. L., GARGANO, M., MATENTZOGU, N., UNNI, D., BRUSH, M., KEITH, D., CONLIN, T., VASILEVSKY, N., ZHANG, X. A., BALHOFF, J. P., BABB, L., BELLO, S. M., BLAU, H., BRADFORD, Y., ET AL. (2020). The Monarch Initiative in 2019: an integrative data and analytic platform connecting phenotypes to genotypes across species. *Nucleic Acids Research* **48**, D704–D715.
- SIMÕES-COSTA, M. & BRONNER, M. E. (2015). Establishing neural crest identity: a gene regulatory recipe. *Development* **142**, 242–257.
- SINGH, A. P. & NÜSSLEIN-VOLHARD, C. (2015). Zebrafish stripes as a model for vertebrate colour pattern formation. *Current Biology* **25**, R81–R92.
- SLATE, J. (2013). From Beavis to beal color: a simulation study to examine how much Qtl mapping can reveal about the genetic architecture of quantitative traits. *Evolution* **67**, 1251–1262.
- *SMITH, R., HEALY, E., SIDDIQUI, S., FLANAGAN, N., STEIJLEN, P. M., ROSDAHL, I., JACQUES, J. P., ROGERS, S., TURNER, R., JACKSON, I. J., BIRCH-MACHIN, M. A. & REES, J. L. (1998). Melanocortin 1 receptor variants in an Irish population. *The Journal of Investigative Dermatology* **111**, 119–122.
- *SPIEWAK, J. E., BAIN, E. J., LIU, J., KOU, K., STURIALE, S. L., PATTERSON, L. B., DIBA, P., EISEN, J. S., BRAASCH, I., GANZ, J. & PARICHY, D. M. (2018). Evolution of endothelin signaling and diversification of adult pigment pattern in *Danio* fishes. *PLoS Genetics* **14**, e1007538.
- STERN, D. L. & ORGOGOZO, V. (2008). The loci of evolution: how predictable is genetic evolution? *Evolution* **62**, 2155–2177.
- STERN, D. L. & ORGOGOZO, V. (2009). Is genetic evolution predictable? *Science* **323**, 746–751.
- *SULEM, P., GUDBJARTSSON, D. F., STACEY, S. N., HELGASON, A., RAFNAR, T., JAKOBSDOTTIR, M., STEINBERG, S., GUDJONSSON, S. A., PÁLSSON, A., THORLEIFSSON, G., PÁLSSON, S., SIGURGEIRSSON, B., THORISDOTTIR, K., RAGNARSSON, R., BENEDIKTSDOTTIR, K. R., ET AL. (2008). Two newly identified genetic determinants of pigmentation in Europeans. *Nature Genetics* **40**, 835–837.
- *SULEM, P., GUDBJARTSSON, D. F., STACEY, S. N., HELGASON, A., RAFNAR, T., MAGNUSSON, K. P., MANOLESCU, A., KARASON, A., PÁLSSON, A., THORLEIFSSON, G., JAKOBSDOTTIR, M., STEINBERG, S., PÁLSSON, S., JONASSON, F., SIGURGEIRSSON, B., ET AL. (2007). Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nature Genetics* **39**, 1443–1452.
- SVENSSON, P. A. & WONG, B. B. M. (2011). Carotenoid-based signals in behavioural ecology: a review. *Behaviour* **148**, 131–189.
- *TAKEUCHI, S., SUZUKI, H., YABUCHI, M. & TAKAHASHI, S. (1996). A possible involvement of melanocortin 1-receptor in regulating feather color pigmentation in the chicken. *Biochimica et Biophysica Acta* **1308**, 164–168.
- *TANAKA, J., LEEB, T., RUSHTON, J., FAMULA, T. R., MACK, M., JAGANNATHAN, V., FLURY, C., BACHMANN, I., EBERTH, J., McDONNELL, S. M., PENEDO, M. C. T. & BELLONE, R. R. (2019). Frameshift variant in *MFS12* explains the mushroom coat color dilution in Shetland ponies. *Genes* **10**, 826.
- *TANAVE, A., IMAI, Y. & KOIDE, T. (2019). Nested retrotransposition in the East Asian mouse genome causes the classical nonagouti mutation. *Communications Biology* **2**, 283.
- *THERON, E., HAWKINS, K., BERMINGHAM, E., RICKLEFS, R. E. & MUNDY, N. I. (2001). The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanistic plumage morph of the bananaquit, *Coereba flaveola*. *Current Biology* **11**, 550–557.
- *TOBITA-TERAMOTO, T., JANG, G. Y., KINO, K., SALTER, D. W., BRUMBAUGH, J. & AKIYAMA, T. (2000). Autosomal albino chicken mutation (*ca/ca*) deletes hexanucleotide ($-\Delta GACTGG817$) at a copper-binding site of the tyrosinase gene. *Poultry Science* **79**, 46–50.
- TOEWS, D. P. L., HOFMEISTER, N. R. & TAYLOR, S. A. (2017). The evolution and genetics of carotenoid processing in animals. *Trends in Genetics* **33**, 171–182.
- *TOEWS, D. P. L., TAYLOR, S. A., VALLENDER, R., BRELSFORD, A., BUTCHER, B. G., MESSER, P. W. & LOVETTE, I. J. (2016). Plumage genes and little else distinguish the genomes of hybridizing warblers. *Current Biology* **26**, 2313–2318.
- *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. & CARNEIRO, M. (2017). High-density lipoprotein receptor *SCARB1* is required for

- carotenoid coloration in birds. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 5219–5224.
- *TOOMEY, M. B., MARQUES, C. I., ANDRADE, P., ARAÚJO, P. M., SABATINO, S., GAZDA, M. A., AFONSO, S., LOPES, R. J., CORBO, J. C. & CARNEIRO, M. (2018). A non-coding region near Follistatin controls head colour polymorphism in the Gouldian finch. *Proceedings. Biological Sciences* **285**, 20181788.
- *UTZERI, V. J., BERTOLINI, F., RIBANI, A., SCHIAVO, G., DALL'OLIO, S. & FONTANESI, L. (2016). The albinism of the feral Asinara white donkeys (*Equus asinus*) is determined by a missense mutation in a highly conserved position of the tyrosinase (TYR) gene deduced protein. *Animal Genetics* **47**, 120–124.
- *UTZERI, V. J., RIBANI, A. & FONTANESI, L. (2014). A premature stop codon in the TYRP1 gene is associated with brown coat colour in the European rabbit (*Oryctolagus cuniculus*). *Animal Genetics* **45**, 600–603.
- *UY, J. A. C., COOPER, E. A., CUTIE, S., CONCANNON, M. R., POELSTRA, J. W., MOYLE, R. G. & FILARDI, C. E. (2016). Mutations in different pigmentation genes are associated with parallel melanism in Island flycatchers. *Proceedings of the Royal Society B: Biological Sciences* **283**, 20160731.
- *VAEZ, M., FOLLETT, S. A., BED'HOM, B., GOURICHON, D., TIXIER-BOICHARD, M. & BURKE, T. (2008). A single point-mutation within the melanophilin gene causes the lavender plumage colour dilution phenotype in the chicken. *BMC Genetics* **9**, 7.
- *VÅGE, D. I., FUGLEI, E., SNIPSTAD, K., BEHEIM, J., LANDSEM, V. M. & KLUNGLAND, H. (2005). Two cysteine substitutions in the MC1R generate the blue variant of the Arctic fox (*Alopex lagopus*) and prevent expression of the white winter coat. *Peptides* **26**, 1814–1817.
- *VÅGE, D. I., KLUNGLAND, H., LU, D. & CONE, R. D. (1999). Molecular and pharmacological characterization of dominant black coat color in sheep. *Mammalian Genome* **10**, 39–43.
- *VÅGE, D. I., LU, D., KLUNGLAND, H., LIEN, S., ADALSTEINSSON, S. & CONE, R. D. (1997). A non-epistatic interaction of agouti and extension in the fox, *Vulpes vulpes*. *Nature Genetics* **15**, 311–315.
- *VÅGE, D. I., NIEMINEN, M., ANDERSON, D. G. & RØED, K. H. (2014). Two missense mutations in melanocortin 1 receptor (MC1R) are strongly associated with dark ventral coat color in reindeer (*Rangifer tarandus*). *Animal Genetics* **45**, 750–753.
- *VAN BUREN, S. L., MINOR, K. M., GRAHN, R. A., MICKELSON, J. R., GRAHN, J. C., MALVICK, J., COLANGELO, J. R., MUELLER, E., KUEHNLEIN, P. & KEHL, A. (2020). A third MLPH variant causing coat color dilution in dogs. *Genes* **11**, 639.
- *VICKREY, A. I., BRUDERS, R., KRONENBERG, Z., MACKAY, E., BOHLENDER, R. J., MACLARY, E. T., MAYNEZ, R., OSBORNE, E. J., JOHNSON, K. P., HUFF, C. D., YANDELL, M. & SHAPIRO, M. D. (2018). Introgression of regulatory alleles and a missense coding mutation drive plumage pattern diversity in the rock pigeon. *eLife* **7**, e34803.
- *VIDAL, O., ARAGUAS, R. M., FERNÁNDEZ, E., HERAS, S., SANZ, N. & PLA, C. (2010a). Melanism in Guinea fowl (*Nimida meleagris*) is associated with a deletion of Phenylalanine-256 in the MC1R gene. *Animal Genetics* **41**, 656–658.
- *VIDAL, O., VIÑAS, J. & PLA, C. (2010b). Variability of the melanocortin 1 receptor (MC1R) gene explains the segregation of the bronze locus in Turkey (*Meleagris gallopavo*). *Poultry Science* **89**, 1599–1602.
- *WADA, A., KUNIEDA, T., NISHIMURA, M., KAKIZOE-ISHIDA, Y., WATANABE, N., OHKAWA, K. & TSUDZUKI, M. (2005). A nucleotide substitution responsible for the tawny coat color mutation carried by the MSKR inbred strain of mice. *The Journal of Heredity* **96**, 145–149.
- *WANG, L., SUN, F., WAN, Z. Y., YE, B., WEN, Y., LIU, H., YANG, Z., PANG, H., MENG, Z., FAN, B., ALFIKO, Y., SHEN, Y., BAI, B., LEE, M. S. Q., PIFERRER, F., ET AL. (2021). Genomic basis of striking fin shapes and colors in the fighting fish. *Molecular Biology and Evolution* **38**, 3383–3396.
- *WANG, Z., QU, L., YAO, J., YANG, X., LI, G., ZHANG, Y., LI, J., WANG, X., BAI, J., XU, G., DENG, X., YANG, N. & WU, C. (2013). An EAV-HP insertion in 5' flanking region of SLC01B3 causes blue eggshell in the chicken. *PLoS Genetics* **9**, e1003183.
- *WIJESENA, H. R. & SCHMUTZ, S. M. (2015). A missense mutation in SLC45A2 is associated with albinism in several small long haired dog breeds. *The Journal of Heredity* **106**, 285–288.
- *WILDE, S., TIMPSON, A., KIRSANOW, K., KAISER, E., KAYSER, M., UNTERLÄNDER, M., HOLLFELDER, N., POTEKHINA, I. D., SCHIER, W., THOMAS, M. G. & BURGER, J. (2014). Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 4832–4837.
- *WINKLER, P. A., GORNIK, K. R., RAMSEY, D. T., DUBIELZIG, R. R., VENTA, P. J., PETERSEN-JONES, S. M. & BARTOE, J. T. (2014). A partial gene deletion of SLC45A2 causes oculocutaneous albinism in Doberman pinscher dogs. *PLoS One* **9**, e92127.
- *WONG, A. K., RUHE, A. L., ROBERTSON, K. R., LOEW, E. R., WILLIAMS, D. C. & NEFF, M. W. (2013). A de novo mutation in KIT causes white spotting in a subpopulation of German Shepherd dogs. *Animal Genetics* **44**, 305–310.
- *WOODCOCK, M. R., VAUGHN-WOLFE, J., ELIAS, A., KUMP, D. K., KENDALL, K. D., TIMOSHEVSKAYA, N., TIMOSHEVSKIY, V., PERRY, D. W., SMITH, J. J., SPIEWAK, J. E., PARICHY, D. M. & VOSS, S. R. (2017). Identification of mutant genes and introgressed tiger salamander DNA in the laboratory axolotl, *Ambystoma mexicanum*. *Scientific Reports* **7**, 6.
- WORTH, C. L., GONG, S. & BLUNDELL, T. L. (2009). Structural and functional constraints in the evolution of protein families. *Nature Reviews Molecular Cell Biology* **10**, 709–720.
- *WRIGHT, H. E., SCHOFIELD, E., MELLERSH, C. S. & BURMEISTER, L. M. (2019). A novel TYRP1 variant is associated with liver and tan coat colour in Lancashire Heelers. *Animal Genetics* **50**, 783.
- *WU, X., ZHANG, Y., SHEN, L., DU, J., LUO, J., LIU, C., PU, Q., YANG, R., LI, X., BAI, L., TANG, G., ZHANG, S. & ZHU, L. (2016). A 6-bp deletion in exon 8 and two mutations in introns of TYRP1 are associated with blond coat color in Liangshan pigs. *Gene* **578**, 132–136.
- *XI, Y., XU, Q., HUANG, Q., MA, S., WANG, Y., HAN, C., ZHANG, R., WANG, J., LIU, H. & LI, L. (2021). Genome-wide association analysis reveals that EDNRB2 causes a dose-dependent loss of pigmentation in ducks. *BMC Genomics* **22**, 381.
- *XU, X., DONG, G.-X., HU, X.-S., MIAO, L., ZHANG, X.-L., ZHANG, D.-L., YANG, H.-D., ZHANG, T.-Y., ZOU, Z.-T., ZHANG, T.-T., ZHUANG, Y., BHAK, J., CHO, Y. S., DAI, W.-T., JIANG, T.-J., ET AL. (2013). The genetic basis of white tigers. *Current Biology* **23**, 1031–1035.
- *XU, X., DONG, G.-X., SCHMIDT-KÜNTZEL, A., ZHANG, X.-L., ZHUANG, Y., FANG, R., SUN, X., HU, X.-S., ZHANG, T.-Y., YANG, H.-D., ZHANG, D.-L., MARKER, L., JIANG, Z.-F., LI, R. & LUO, S.-J. (2017). The genetics of tiger pelage color variations. *Cell Research* **27**, 954–957.
- *YAN, S. Q., HOU, J. N., BAI, C. Y., JIANG, Y., ZHANG, X. J., REN, H. L., SUN, B. X., ZHAO, Z. H. & SUN, J. H. (2014). A base substitution in the donor site of intron 12 of KIT gene is responsible for the dominant white coat colour of blue fox (*Alopex lagopus*). *Animal Genetics* **45**, 293–296.
- *YANG, J., CHEN, X., BAI, J., FANG, D., QIU, Y., JIANG, W., YUAN, H., BIAN, C., LU, J., HE, S., PAN, X., ZHANG, Y., WANG, X., YOU, X., WANG, Y., ET AL. (2016). The *Sinoocycheilus* cavefish genome provides insights into cave adaptation. *BMC Biology* **14**, 1.
- *YU, F., JIAO, S., LAI, W., LIU, Z., ZHU, M., ZHU, W., BAI, C., ZHANG, Y., ZHANG, J. & YAN, S. (2018). Conserved aspartate-to-glycine mutation in tyrosinase is associated with albino phenotype in domestic Guinea pigs (*Cavia porcellus*). *Animal Genetics* **49**, 354–355.
- *YU, S., WANG, G. & LIAO, J. (2019a). Association of a novel SNP in the ASIP gene with skin color in black-bone chicken. *Animal Genetics* **50**, 283–286.
- *YU, W., WANG, C., XIN, Q., LI, S., FENG, Y., PENG, X. & GONG, Y. (2013). Non-synonymous SNPs in MC1R gene are associated with the extended black variant in domestic ducks (*Anas platyrhynchos*). *Animal Genetics* **44**, 214–216.
- *YU, Y., GRAHN, R. A. & LYONS, L. A. (2019b). Mocha tyrosinase variant: a new flavour of cat coat coloration. *Animal Genetics* **50**, 182–186.
- *YUSNIZAR, Y., WILBE, M., HERLINO, A. O., SUMANTRI, C., NOOR, R. R., BOEDIONO, A., ANDERSSON, L. & ANDERSSON, G. (2015). Microphthalmia-associated transcription factor mutations are associated with white-spotted coat color in swamp buffalo. *Animal Genetics* **46**, 676–682.
- *ZHANG, M.-Q., XU, X. & LUO, S.-J. (2014). The genetics of brown coat color and white spotting in domestic yaks (*Bos grunniens*). *Animal Genetics* **45**, 652–659.
- *ZHOU, Z., LI, M., CHENG, H., FAN, W., YUAN, Z., GAO, Q., XU, Y., GUO, Z., ZHANG, Y., HU, J., LIU, H., LIU, D., CHEN, W., ZHENG, Z., JIANG, Y., ET AL. (2018). Author Correction: An intercross population study reveals genes associated with body size and plumage color in ducks. *Nature Communications* **9**, 3974.

VIII. SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

File S1. Data set used in this analysis downloaded from *Gephebase* on 15 August 2022 (tab: All Vertebrates), the gene ontology (GO) assignments for each gene (tab: GO Assignments), and the final molecular function categories including child GO terms (tab: Molecular Function Summary).

Fig. S1. Relative proportions of *cis*-regulatory and coding mutations associated with each study methodology, and in the total data set, after correcting for the number of entries associated with each gene.

Fig. S2. Relative proportions of *cis*-regulatory and coding mutations associated with different taxonomic status (domesticated, intraspecific and interspecific variation) and study methodology, and the total proportion for each taxonomic status.

Fig. S3. (A) Relative proportions of *cis*-regulatory and coding mutations associated with each taxonomic status and clade, and in the total data set, after correcting for the number of entries associated with each gene.

Fig. S4. Relative proportions of *cis*-regulatory and coding mutations associated with different taxonomic status (domesticated, intraspecific variation, and interspecific variation) and clade, and the total proportion for each taxonomic status, after removing ambiguous cases.

Fig. S5. Relative proportions of *cis*-regulatory and coding mutations associated with different phenotype categories

according to taxonomic status: domesticated, intraspecific, and interspecific, and the total proportion for each taxonomic status, after removing ambiguous cases.

Fig. S6. Relative proportions of *cis*-regulatory and coding mutations associated with different phenotype categories according to study methodologies: candidate gene, linkage mapping, and association mapping, and the total proportion for each study methodology, after removing ambiguous cases.

Fig. S7. (A) Relative proportions of *cis*-regulatory and coding mutations associated with each phenotype category, functional category, and molecular function category, and the total for the data set, after correcting for the number of entries associated with each unique gene.

Table S1. The 10 genes for each clade with the most total entries across other clades, excluding amphibians.

(Received 2 February 2022; revised 8 March 2023; accepted 14 March 2023; published online 5 April 2023)