PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Review

Cite this article: San-Jose LM, Roulin A. 2017 Genomics of coloration in natural animal populations. *Phil. Trans. R. Soc. B* **372**: 20160337.
http://dx.doi.org/10.1098/rstb.2016.0337

Accepted: 15 February 2017

One contribution of 19 to a theme issue 'Animal coloration: production, perception, function and application'.

Subject Areas:
genomics, evolution, ecology

Keywords:
animal coloration, genomic-wide association studies, next-generation sequencing, RNA-Seq, colour evolution

Author for correspondence:
Luis M. San-Jose
e-mail: luis.sanjosegarcia@unil.ch

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3732259.

THE ROYAL SOCIETY PUBLISHING

Genomics of coloration in natural animal populations

Luis M. San-Jose and Alexandre Roulin

Department of Ecology and Evolution, University of Lausanne, Building Le Biophore, 1015 Lausanne, Switzerland

LMS, 0000-0003-1351-4679; AR, 0000-0003-1940-6927

Animal coloration has traditionally been the target of genetic and evolutionary studies. However, until very recently, the study of the genetic basis of animal coloration has been mainly restricted to model species, whereas research on non-model species has been either neglected or mainly based on candidate approaches, and thereby limited by the knowledge obtained in model species. Recent high-throughput sequencing technologies allow us to overcome previous limitations, and open new avenues to study the genetic basis of animal coloration in a broader number of species and colour traits, and to address the general relevance of different genetic structures and their implications for the evolution of colour. In this review, we highlight aspects where genome-wide studies could be of major utility to fill in the gaps in our understanding of the biology and evolution of animal coloration. The new genomic approaches have been promptly adopted to study animal coloration although substantial work is still needed to consider a larger range of species and colour traits, such as those exhibiting continuous variation or based on reflective structures. We argue that a robust advancement in the study of animal coloration will also require large efforts to validate the functional role of the genes and variants discovered using genome-wide tools.

This article is part of the themed issue 'Animal coloration: production, perception, function and application'.

1. Past and future of the genetics of animal coloration in natural populations

The study of animal coloration has been essential for the development of biological sciences, particularly for the fields of genetics and evolutionary biology (reviewed in [1]). Pioneering geneticists like Morgan, Bateson or Haldane studied the inheritance of colour traits to establish the basics of Mendelian genetics. Similarly, studies of coloration in wild populations, such as Kettlewell's studies [2] on the action of natural selection on the black and pale morphs of the peppered moth, *Biston betularia*, or Endler's studies [3] on the joint role of sexual and natural selection on explaining colour variation in guppies, *Poecilia reticulata*, have also moulded our understanding of evolution. For practical reasons, many of these studies took advantage of colour traits exhibiting relatively simple discrete variation and inheritance patterns. Also for practical reasons, research on animal coloration has continued and flourished mainly on the model systems used by these pioneers (*Drosophila* and mice, for instance), with only a few new species being adopted as model systems later on (e.g. the zebrafish, *Danio rerio* [4]). Our current understanding of the genetics of animal coloration is therefore limited to a handful of well-studied systems, and does not account for the extant variation of life forms and colour traits. In natural conditions, discrete variation is the exception rather than the rule. Most colour traits vary continuously between two extreme values and are complexly based on the deposition of several pigments and/or on the spatial arrangement of integumentary structures. The expression of these traits can be strongly determined by genetic factors but also by the environment or by the interaction of genetic and environmental factors. Finally, colour traits can have different types of functions and adaptive roles

[5–8], which can result in different evolutionary histories and underlying genetic architectures, even when considering colour traits of high resemblance.

Remarkable exceptions aside (e.g. mapping studies in *Heliconus*, cichlids and white-throated sparrows, *Zonotrichia albicollis* [5]), identifying the underlying genes and genetic structures of such a diverse spectrum of colour traits has received little attention in the wild, partly because previous methods to find loci responsible for colour variation were laborious, expensive and/or unfeasible in natural populations. In recent years and thanks to the seminal work of N. Mundy [9] and H. Hoekstra [10] on the melanocortin-receptor 1, *MC1R*, unravelling the genetic basis of coloration has been mainly based on the more accessible candidate-gene approach [5,8]. This method assesses the role that genes known to regulate coloration in model species play in non-model species. It has proven very useful to achieve a better understanding of colour variation in a wide range of species, highlighting the recurrent role of certain colour genes, like the *MC1R*, in mediating adaptive colour changes in several species [5]. However, the candidate-gene approach is severely limited given that candidate genes are described only in the few existing model species, and on the basis of their role in the colour traits these species display. As a consequence, the study of the genetic architecture of colour traits has been biased towards melanin-based coloration, the main source of colour variation in current model species such as mice and *Drosophila*. Additionally, the discovery of new genes and the study of colour traits other than melanin are hindered, if not impossible, with the candidate gene approach.

The study of animal coloration has often neglected the role of the genetic structure underlying colour variation. Explicitly or implicitly, phenotypic variance has been assumed to be a good surrogate of genetic variance, and phenotypic evolution to circumvent the constraints imposed by the genetic architecture. This assumption, the so-called 'phenotypic gambit' [11], has allowed us to simplify evolutionary scenarios and to model and test the applicability of theories that can explain how colour traits evolve. For instance, the burst of studies conducted on carotenoid-based coloration in the previous decade was largely influenced by the modelling that Grafen and others did of the Zahavi's handicap principle [12]. This provided a good theoretical background to understand how carotenoid-based coloration could have evolved, and to envision the evolutionary consequences of observations done at the phenotypic level. However, the phenotypic gambit has rarely been challenged in relation to colour traits. To the best of our knowledge, the only study conducted in this sense rejected an association between phenotypic and genetic variation for colour traits [13], highlighting the need for improving our knowledge on the genetic basis of colour variation in order to understand how colour forms evolve.

The arrival of the 'genomic era' with high-throughput or next-generation sequencing methods opens new avenues for the study of the genetics of animal coloration in wild populations, surpassing limitations of previous approaches to a great extent. Current genomic tools (reviewed in [14]) offer the opportunity to conduct genetic studies with a finer detail, incurring fewer costs than a few decades ago. They do not entirely rely on previous knowledge of particular genes, and they can be applied to outbred populations and model systems where conducting controlled crosses or obtaining pedigree information is impractical. Consequently, genomic tools allow us to explore genetic aspects that were only possible to address in model species. This widens the spectrum of colour traits, species, and ecological and evolutionary scenarios to study, filling current knowledge gaps and leading to more robust generalizations on the evolution, development and ecology of animal colorations. These new sequencing tools have been promptly adopted to study natural variation of coloration in non-model species (figure 1 and table 1). Genomic studies conducted to date depict a diverse spectrum of genetic architectures, including single coding mutations underlying colour variation between populations, alternative *cis*-regulatory changes controlling colour variation within and between species, and supergenes associated with discrete colour morphotypes, among others (figure 1 and table 1). However, biases towards the study of melanin-based colorations, discrete colour traits and specific groups (e.g. mimicry in butterflies and colour patterns in cichlid fish, table 1) still prevail. We believe that several knowledge gaps still remain and that it is timely to highlight them as well as how high-throughput sequencing can be put into use to improve our understanding of the underlying genetics of colour diversity in animals.

Here, we review how different types of colorations are generated, how colorations within and between body parts are genetically integrated, what genetic mechanisms drive associations between colour and other phenotypes, and some of the genetic assumptions yielded by theoretical models of colour evolution. Research in other important aspects of animal colour evolution (e.g. coevolution of coloration and colour vision, the role of animal coloration in reproductive isolation and speciation, the genetic basis for rapid morphological colour changes, among others) can also benefit from the application of high-throughput sequencing data. Unfortunately, particular consideration of all aspects regarding animal coloration is beyond the scope of this review. However, many of the principles reviewed here have general implications for our understanding of the basics of animal coloration and are therefore expected to be also of use for a broader range of research on animal coloration.

2. Complicating the picture: beyond melanin and single-trait variation

Colour production can be based on different pigments (e.g. melanins, pteridines, carotenoids, ommochromes, porphyrins), or reflective structures deriving from the finely tuned arrangement of purine crystals, collagen, keratin or chitin [26]. The synthesis pathways are known for most of these components (e.g. [27]), except for taxa-specific pigments such as psittacofulvins and turacoverdins. In contrast, the regulation of other processes responsible for variation in animal coloration (e.g. pigment transport and allocation or the fine spatial organization of reflective structures) is less well understood except perhaps for melanins, for which the regulation of melanosome formation and of melanophores and melanoblasts migration, and differentiation in different vertebrate taxa and the process of cuticular melanin deposition in different invertebrate taxa has been studied [4,28–30]. The genetic basis of melanin-based coloration has also been more extensively studied in the wild, in contrast to other pigments, and several genes such as the *MC1R* and *OCA2* [5] have been associated with variation in melanin-based coloration in natural populations. However, the number of studies where candidate

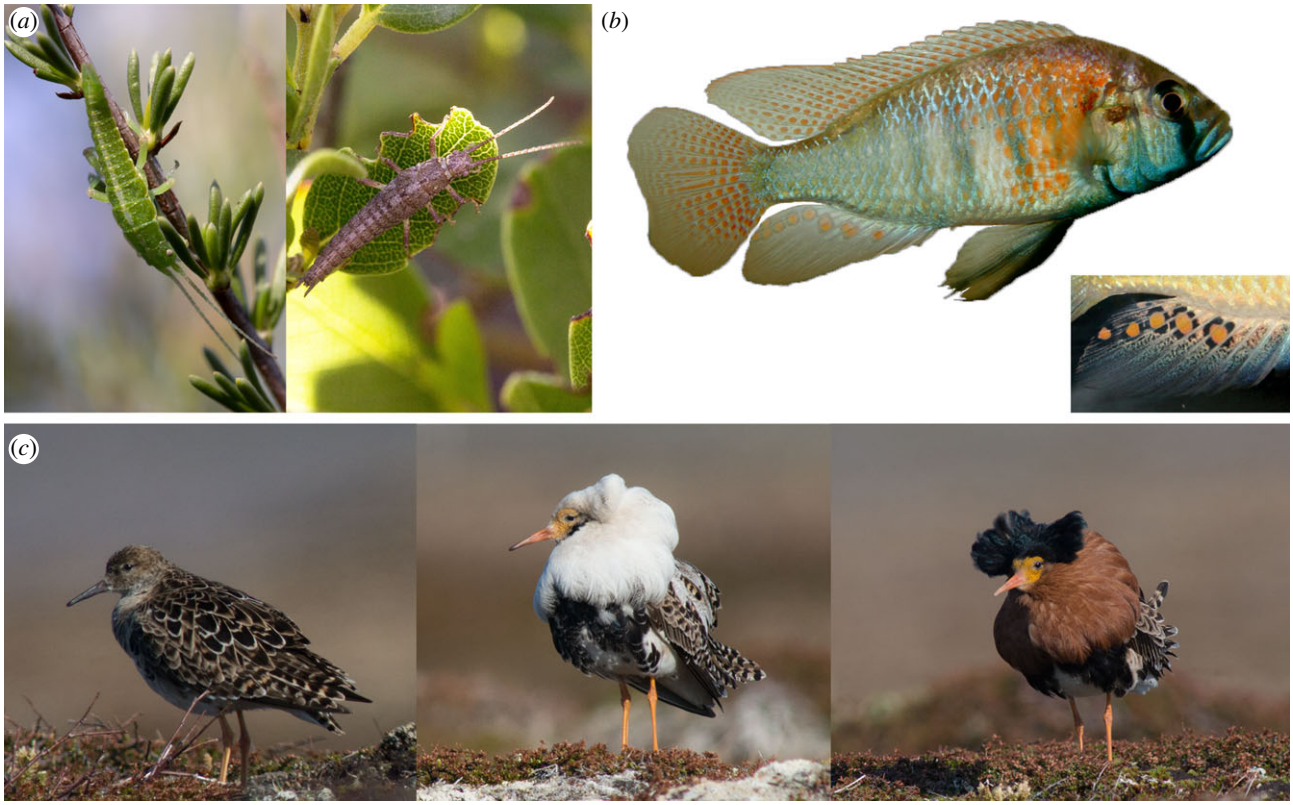

Figure 1. Studying the genetic basis of coloration in natural populations using high-throughput sequencing. The figure illustrates three case studies where the genetic basis of colour variation has been studied using high-throughput sequencing methods. (*a*) Using genetic crosses and genome-wide association mapping based on RAD-sequencing markers, the genetic architecture of green/brown polymorphism and pattern (i.e. the presence/absence of the light longitudinal stripe) was unravelled in *Timema cristinae* (see [15,111,112] in table 1). One and two large-effect loci were found for colour and pattern, respectively. Dominant effects of the alleles determining green and stripe occurrence were observed for all loci, and such genetic architecture might be constraining local adaptation in the green form. (*b*) Using whole-transcriptome shotgun sequencing together with several validation methods (see text and ref. [16]), the role of the gene *fhl2b* in mediating egg-spot formation in the anal fins of haplochromine cichlids was demonstrated. The haplochromine cichlid *Astatotilapia burtoni* and a detailed image of the anal fin are shown. (*c*) Using a combination of RAD-sequencing and whole-genome resequencing, a genomic region hosting ca. 125 genes was found to differ between the three forms (faeder, satellite and independent) present in male ruffs, *Philomachus pugnax*. Recombination between forms is suppressed owing to two alternative and recessive lethal genomic inversions of this region in faeder and satellite forms (see [17]). Pictures were kindly provided by A. Comeault and P. Nosil (*a*), E. Santos and A. Theis (*b*) and R. Vervoort (*c*).

genes have failed to explain colour variation is accumulating [31], indicating that there must be other genetic actors driving variation in melanin-based coloration in natural populations.

Despite some recent discoveries [19,20] and advancement in understanding the biology of xanthophores in model species [4], carotenoid-based colorations still have an understudied genetic basis. To the best of our knowledge, the genetic basis of coloration based on pteridines has not been studied in natural populations, despite having been intensively studied in *Drosophila* eye mutants and zebrafish [4], and being a common pigment in both vertebrates and invertebrates. Ommochromes have received relatively more attention, given that they are responsible for colour variation in *Heliconius* butterflies, where genes like *optix* were shown to regulate the development of red wing patterning in several species [32]. In relation to structural components, genes that were highlighted in screenings of zebrafish mutants have also been observed to underlie iridophore-based coloration in a natural cichlid population [33], and Santos *et al*. identified two novel candidate genes regulating iridophore-based coloration in haplochromine cichlids [16]. However, little is known about the regulation of other structural components [34] or the regulation of iridophore-based coloration outside fish.

Only rarely does animal coloration result from a single pigmentary or structural component, and a deeper understanding of how colour differences arise also comes from studying how different colour components interact [35–37]. In the same sense, the development and function of a given colour patch has been shown to depend on coloration at other body parts (i.e. on colour patterning), which can also be based on different pigments or structural components. This is best exemplified by recent studies conducted on the striped pattern of adult zebrafish [4]. In the dark stripes, xanthophores are faintly pigmented, stellated and have lower densities than in the light stripes. Such differences in xanthophore morphology and number result from interactions with the other types of skin chromatophores: iridophores and melanophores. Different connexins mediate these interactions, and mutations at their encoding genes do not only impact xanthophore differentiation but the formation of the whole pattern, with dark stripes dissolving into dark spots. Spatial and temporal modularization of gene expression via transcription factors or epigenetic changes is also expected to be of great importance to differently use the same genetic machinery at distinct body parts [28]. This aspect has been studied in some natural scenarios, for instance in beach mice (*Peromyscus polionotus*), where differential *cis*-regulation of *ASIP* transcription seems to account for differences in colour patterning between populations [38], or in carrion crows (*Corvus corone*), where differential expression of *ASIP* across body parts was observed even in uniformly

Table 1. Studies conducted on animal coloration using high-throughput (next-generation) sequencing methods. WGS: whole-genome sequencing. RAD-seq: Restriction-site-associated sequencing. RNA-seq: transcriptome shotgun sequencing.

	colour trait	variation	level of variation	genomic data	main finding/s	reference[a]
Aves						
ruff (*Philomachus pugnax*)	melanin	discrete	within population	RAD-seq, WGS	Inversion at chromosome 11 containing 125 genes including the *MC1R* gene.	[17]
ruff (*Philomachus pugnax*)	melanin	discrete	within population	RNA-seq	No genes were found to be significantly differently expressed between male morphs. Several SNP markers, many in genes of unknown origin, were found associated with male colour morphs.	[88]
carrion and hooded crow (*Corvus corone corone, C. c. cornix*)	melanin	discrete	between subspecies	WGS, RNA-seq	A single narrow genomic region was identified to be divergent and to include genes known to be associated with melanogenesis-related transcription factor *MITF*. Several melanogenic-related genes, but not *MITF*, were observed to be differentially expressed in the feathers of both subspecies.	[89]
carrion and hooded crow (*Corvus corone corone, C. c. cornix*)	melanin	discrete	within individual and between subspecies	RNA-seq	Several genes involved in melanogenesis and regulation of melanogenesis were found to be differentially expressed between subspecies.	[18]
red siskins (*Spinus cucullata*)	carotenoid	discrete	between species	WGS	By crossing red siskins with common canaries, Lopes *et al*. found a region potentially responsible for red coloration in siskins. This region included the gene *CYP2J19*, which was also found to be more expressed. In parallel, Mundy *et al*. mapped the *yellowbeak* mutation in zebra finches to a region also containing *CYP2J19* and two other related genes and observed differential gene expression of some of these genes in wild and mutant zebra finches.	[19,20]
golden- and blue-winged warblers (*Vermivora chrysoptera, V. cyanoptera*)	melanin, carotenoid	discrete	between species	WGS	Six divergent regions between species were found, four that include genes known to be associated with feather development and melanin-based colorations (*ASIP*, *BCO2*, *Wnt*).	[90]
white-throated sparrow (*Zonotrichia albicollis*)	melanin	discrete	within population	WGS	Characterization of the previously described 'supergene', which includes approximately 1000 genes with several potential candidates to explain colour differences between morphs.	[91]
several species	melanin	discrete	between species	WGS	Evidence for coevolution of colour (*MC1R*) and opsins (*RH2*) genes, which encode for visual pigments.	[92]

(*Continued*.)

Table 1. (*Continued.*)

	colour trait	variation	level of variation	genomic data	main finding/s	reference[a]
Holarctic redpoll finches (*Acanthis spp.*)	melanin	continuous	between species	RAD-seq, RNA-seq	Evidence for scarce genomic divergence between species and for gene expression differences in association differences in plumage and bill morphology.	[93]
flycatchers (*Z. chrysops, Z. viridiflavus*)	melanin	discrete	between species	RAD-seq	Evidence for introgression as a driver of phenotypic similarities in plumage traits. Introgressed regions were enriched for genes involved in 'cellular components'.	[94]
several species	mixed	—	between sexes	RAD-seq	Evidence for sex-linked genetic variation being smaller in monochromatic than in dichromatic bird species.	[95]
buzzards (*Buteo swainsoni, B. jamaicensis, B. lagopus*)	melanin	discrete	within-species	RAD-seq	Several SNPs (no annotation provided) were found in association with variation in coloration.	[96]
chestnut-bellied monarch (*Monarcha castaneiventris*)	melanin	discrete	within and between populations	RAD-seq	Confirmation of two candidate SNPs at *MC1R* and *ASIP* involved in colour differences while controlling for population structure.	[97]
several species	mixed	discrete	between species	sequence capture	Phylogenetic reconstruction and comparative analysis suggests multiple gains and losses of eyespot evolution in *Galliformes*.	[98]
common buzzard (*Buteo buteo*)	melanin	discrete	within individual	RNAseq	identification of 28 candidate genes associated with dorsal-ventral differences in coloration.	[99]
collared flycatcher (*Ficedula albicollis*)	structural	continuous	within population	WGS	No candidate loci were found but detailed power analyses are provided.	[21]
woodpeckers (*Sphyrapicus ruber, S. nuchalis, S. varius*)	mixed	discrete	between species	RAD-seq	One SNP marker at *COG4* was found to be associated with colour variation.	[100]
mammals						
oldfield mouse (*Peromyscus polionotus*)	melanin	discrete	between subspecies	sequence capture	Evidence for standing genetic variation at the *MC1R* in mainland populations leading to parallel evolution of apigmented coastal populations. No evidence for genomic patterns of selection in the surroundings of the *MC1R*.	[101]
wolf (*Canis lupus*)	melanin	discrete	within species	sequence capture	Eight candidate SNPs at *CBD103* were found to be associated with coat colour differences.	[102]
reptiles						
several species	melanin	discrete	between and within species	sequence capture	Evidence for selection at the *MC1R* gene in 'blanched' populations.	[103]

(*Continued.*)

Table 1. (*Continued.*)

	colour trait	variation	level of variation	genomic data	main finding/s	reference[a]
fish						
Astatotilapia burtoni	mixed	discrete	within individual	RNA-seq	Many (approx. 1000) genes were found to be differentially expressed in association to egg-spots.	[104]
Astatotilapia burtoni	mixed	discrete	between sexes, between species	RNA-seq	Implication of *fhl2* paralogs in iridophore-based coloration and evidence for evolution of *fhl2b* regulation by the insertion of a transposable element.	[16]
Sardine cichlid (*Cyprichromis leptosoma*)	mixed	discrete	within species	RAD-seq	Evidence for a bi-allelic QTL controlling male polymorphism and for strong linkage in the surroundings of that locus suggesting an inversion.	[105]
Metriaclima zebra, M. mbenji	mixed	discrete	between species	RAD-seq	QTLs for melanophore and xanthophore variation were identified and several candidate genes (*Csfr1*, *TSPM* 6/1) in the proximities of the QTLs were proposed.	[106]
three-spined stickleback (*Gasterosteus aculeatus*)	carotenoid	discrete	within population	RAD-seq	Several genomic regions were found to be highly differentiated between different colour morphs. These regions did not overlap with previously described candidate colour QTLs for the species.	[107]
Midas cichid (*Amphilophus citrinellus*)	mixed	discrete	within individual	RNA-seq	Identification of 46 candidate genes showing differential expression between age-related morphs.	[108]
Haplochromis sauvagei, H. nyererei	melanin	discrete	between species	RAD-seq	Evidence for a genomic region containing the gene *TBLX1* associated with the transmission of lateral stripes.	[109]
Insects						
stick insects (*Timema cristinae, T. podura*)	melanin	discrete	within population, between species	RAD-seq	Dominant, major-effect loci associated colour were found in both species and allocated to the same genomic region (linkage group).	[15,110,111]
Heliconius butterflies	mixed	discrete	between species	RAD-seq	Evidence for hybrid exchange of genes related to mimicry patterns.	[112]
Heliconious butterflies	mixed	discrete	between species	sequence capture	Evidence for the *cortex* gene controlling patterning in several species.	[22]
Papilio polytes	mixed	discrete	between subspecies	RAD-seq, WGS	A single gene, *doublesex*, was found associated with sex-limited mimicry.	[113]
Heliconius and *Limenitis* butterflies	mixed	discrete	between and within species	WGS, RNA-seq	Candidate SNPs upstream of the coding region of the *WntA* gene were found in both species.	[23]
Heliconius hecale, H. ismenius	mixed	discrete	between and within species	RAD-seq	Several genomic regions known to associate to colour differences in other *Heliconus* species were found.	[114]

(*Continued.*)

Table 1. (*Continued.*)

	colour trait	variation	level of variation	genomic data	main finding/s	reference[a]
Papilio xuthus, P. machaon	mixed	discrete	between species	WGS, RNA-seq	Evidence for different divergent regions, exhibiting differential expression and selective pressures.	[24]
Vanessa cardui	mixed	discrete	within individual	RNA-seq	Identification of genes differentially expressed during the formation of wing colour pattern. The role of three candidate genes (*Abdominal-B*, *ebony* and *frizzled*) within the differentiated regions were validated using CRISPR/Cas9.	[115]
Biston betularia	melanin	discrete	within population	sequence capture	Evidence for an insertion of a transposable element in *cortex* as responsible for melanism.	[116]
Ischnura elegans	mixed	discrete	within individual	RNA-seq	Identification of genes differentially expressed between differently coloured body parts.	[117]
Bombus bifarius	mixed	discrete	between populations	RNA-seq, RAD-seq	Population genomics of geographically distant morphs.	[118]
Drosophila melanogaster	melanin	continuous	within, between populations	WGS	Several candidate SNPs found in *tan* and *bab1to* to associate with colour variation in different populations although with different relative importance in each population.	[119]
Drosophila melanogaster	melanin	continuous	within population	sequence capture	Several candidate SNPs were identify clustering in different regulatory regions of the *bab* locus.	[120]
Drosophila melanogaster	melanin	continuous	within population	WGS	Several candidate SNPs at *tan*, *ebony* and *bab* where identify to associate to coloration.	[25]
mollusc						
Cepaea nemoralis	melanin	discrete	intra-species	RAD-seq	Evidence for 44 candidate markers linked to differences in colour and banding.	[121]

[a]References above 87 are included in the electronic supplementary material.

black-coloured crows [18]. This points to the existence of interactions with *trans*-acting factors in determining pattern differences between populations in this species. In contrast, epigenetic changes have received little attention in wild populations [39] although they are known to underlie certain colour patterns in domestic or laboratory animals [40].

A deeper knowledge of the genetic architecture of different types of pigments and reflective structures, as well as a more complete understanding of the factors driving their interactions within and across body parts, is still needed. This will help us to understand several unsolved questions, such as the predominance of certain pigments (e.g. melanin) across the tree of life, the existence of different pigments or structures with equivalent effects on coloration (e.g. carotenoids and pteridines), the evolution of taxon-specific components, the loss and gain of certain components in different taxa (e.g. why penguins are the sole known bird taxon colouring their feathers with pteridines), and to what extent the same type of colour traits evolved similar or different genetic architectures under different evolutionary scenarios. Studies in natural conditions will give insight into whether complex colour patterns can indeed evolve from relatively simple genetic mechanisms as depicted by studies in zebrafish, to what extent the pathways of different pigments and/or structures have commonalities, and whether such common pathways could explain the fact that different types of colorations have been repeatedly observed to associate with the same phenotypes (e.g. competitive dominance in intraspecific contexts [41]).

Studies using high-throughput sequencing methods would shed some light on the genetic basis of different types of pigments and structures, the genetic causes of their variation between and within species, and on the genetic basis of their interaction in determining colorations within and between body parts. Whole-transcriptome sequencing tools can be of great utility to unravel the expression of which genes underlie the development of different types of coloration in different taxa and in different body parts (table 1). The particular effects of multiple colour components on a single colour trait could also be addressed by applying more precise methods, such as single-cell RNA sequencing [42]. In combination with whole-genome or genome-wide sequencing techniques, these studies can potentially reveal the mutational changes leading to alternative colours and patterns [22,23]. Furthermore, epigenomics can help us to understand the basis for the modification of gene expression and differential colour expression between and within individuals [43].

Probably, the main limitation to the application of these methods to a wider range of non-model species, colour traits and patterns is the availability of a good reference genome sequence. However, the genomes of several non-model species have recently been assembled or are being assembled (genome10 k.soe.ucsc.edu/), and the arrival of the third-generation sequencing technologies is expected to facilitate and speed up the process of sequencing and assembly of new genomes by producing longer fragments (approx. 10–30 kb). Despite its costs, investing in a reference genome is always desirable. It allows us to conduct more detailed (e.g. whole-genome resequencing: [44]) and accurate [45] analyses given that less information is lost when mapping the sequenced reads [46]. Nevertheless, if costs or time only allow producing a low-quality genome, it can be more advisable to use the genome of a closely related species, or reference-free methods like *de novo* transcriptome assembly rather than to base the analyses on an incomplete and highly fragmented genomic sequence [45,47,48]. Actually, given the increasing number of species with an available genome sequence and the costs of a *de novo* genome assembly, we can expect reference-guided genome or transcriptome assembly methods to become more predominant in the near future. Given this, we encourage further research on the impact that divergence between the study and the reference species has on both assembly and downstream analysis (e.g. [45]).

3. The genetic architecture of continuous colour variation

Animal colour traits can vary continuously between two extreme colour phenotypes, and this variation is of interest because it is often the target of natural and/or sexual selection. Both genetic and environmental factors can be responsible for gradual rather than abrupt colour differences between individuals, although primarily environmental influences have been studied in the wild [49,50]. The underlying genetic structure of such quantitative colour traits is expected to be complex, and its characterization challenging [51]. Loci contributing to quantitative traits can be numerous (several dozen) with most of them having small phenotypic effects. Major differences exist in this sense and only a few major-effect loci can explain most of the variation of quantitative traits [52]. This is best exemplified by the genetics of human eye colour for which six variants of the 37 described so far explain most of the variation [53]. Many causal variants can also occur at very low frequencies (less than 5%) which together with their small effect size undermine the statistical power to detect them in a mapping study [54]. Lack of statistical power to detect small-effect loci is the major limitation when studying the genetic basis of quantitative traits. This is evidenced by the low proportion of heritable variation that is normally explained by the sum effects of the loci discovered in a mapping study [55]. Because many loci may be not detected in mapping studies, it is necessary to take into account the power of individual studies when comparing the genetic structures of different traits.

High-throughput sequencing studies can be more powerful because they substantially increase the number of genetic markers and, for the specific case of whole-genome resequencing, because the requirement of strong linkage between a genetic marker and the causal variant is no longer needed (actually, a rapid decay of linkage over distance can be desirable in order to narrow down the number of candidate causal variants, although see ref. [21]). Whole-genome resequencing has been applied to disentangle the genetic basis of natural, continuous variation in the melanin-based pigmentation of *Drosophila melanogaster* females (table 1). Female pigmentation was found to associate to several SNPs at the genes *tan* and *bab1* in both studies [25,56], although in one of the studies, a less restrictive confidence level and posterior functional validation supported the association of 17 additional variants with female pigmentation [25]. In this study, variation explained by top variants (those associated with *tan*, *bab1* and *ebony*) where found to have major phenotypic effects but to explain only one third of colour heritability. Whole-genome resequencing has also been applied to unravel the genetic basis of the forehead white patch size of male

flycatchers, *Ficedula albicollis* [21]. In their study, Kardos *et al*. [21] maximized the sampling variance by comparing the genomes of individuals with extreme colour phenotypes, a more powerful approach than considering all the intermediate colour values in the population. However, no loci were detected owing to a lack of statistical power to detect loci of moderate effect (less than 10% of the total phenotypic variance), and the apparent absence of larger effect loci [21]. This study suggests that detecting loci of moderate effect size would have required approximately 250 individuals. It also suggests that methods that yield a lower density of markers (e.g. RAD-sequencing [57]) than whole-genome resequencing can have reduced power to detect moderate effect-size loci unless they are applied to populations with small effective size, where higher levels of linkage between genetic markers and the causal variants would be expected owing to a lower genetic variation.

Although this last study could discourage researchers from studying the genetic basis of continuous colour variation in wild species, we remain optimistic. On one hand, it is probably impossible to unravel the genetic basis of a quantitative trait in one single study. The main reason is that the power of the study largely depends on *a priori* unknown features (effect size and allele frequencies). Similarly to Kardos *et al*. [21], the first approximation to the question can help us to gather the information needed to conduct an accurate power analysis and use this analysis to decide what could be the best design. Sequencing more individuals *a posteriori* is always possible (although batch effects should be taken into account: [58]), and probably more economically feasible in the years to come. Alternative designs can be used to increase power to detect small effect variants. The power of association studies in humans increases when individuals are selected not only on the basis of their phenotype (e.g. a disease) but on the basis of their 'familiality' (e.g. the presence of the disease on the family) [59]. In wild animal populations, such an approach can be implemented by gathering pedigree information from either behavioural or molecular data, and by using estimated breeding values (an estimate of the capacity of an individual to genetically transmit the phenotype), to decide which individuals could be included in the association study. When possible, crosses of homozygous lines (alternatively, pedigrees [60]) can be used to create a linkage map using methods like RAD-sequencing that are less costly and can be implemented in a larger number of individuals [61]. Highly polygenic architectures may be partially represented in the parental generation of a cross [62], which can be taken into account by using replicated crosses of different individuals. Finally, if continuous colour traits exhibit substantial environmental variation, studies can be designed to account for known sources of environmental noise by manipulating the environment and scanning for QTLs within and across environments, which can also unravel genotype x environment interactions.

4. Colour and genetically correlated traits

Colour traits have often been found to genetically correlate with behavioural, morphological, physiological and life-history traits [7,63]. Selective forces (e.g. [64]) allowing distinct colour forms with alternative suites of phenotypes to coexist within the same population have received relatively more attention than their genetic architecture. Nevertheless, their genetic architecture is also essential to ensure the co-segregation of coloration with other phenotypes, and therefore, the maintenance of trait associations in the face of reproduction and recombination [63]. Pleiotropy [65,66] and the evolution of adaptive gene complexes (supergenes: [63]) are often invoked to explain colour-trait associations. Both mechanisms have received empirical support in a few species [17,63,67], but further case studies are nevertheless needed. On one hand, pleiotropic effects have been observed to mediate associations between coloration and single phenotypes [67], but whether they can also mediate associations with a suite of different phenotypic traits has not been considered in the same study. Similarly, studies that detected the existence of supergenes used colour morphotypes as the variable of interest, but to the best of our knowledge, parallel association studies also considering the phenotypes that differ between colour morphs have not been conducted. This would reveal whether only variants at the supergenes associate to the different phenotypes or if causal variants for these phenotypes also map outside supergenes. Further case studies would also help us to understand the relative importance of these mechanisms, the conditions that determine why some species evolved one mechanism and not the other, to what extent both mechanisms interact (e.g. how often supergenes host pleiotropic genes or alleles; [68]), and the plausibility of alternative unexplored architectures (e.g. those based on epistatic modifiers: [69]). It is important to highlight that current methods are more powerful to detect large, divergent genomic regions, than to detect pleiotropic effects of single variants, which should be taken into account when considering the role of pleiotropy and supergenes in driving associations between colour and other traits.

Genomic tools in combination with pedigree information can be used to develop detailed genetic maps of distinct colour morphs in order to detect differences in linkage and genomic inversions. As recently done in the ruff, *Philomachus pugnax*, [17], these studies can be combined with deep sequencing of the inverted region (or regions) to detect the genes included inside it. Owing to the strong linkage within the non-recombining region, candidate variants need to be tested and validated with posterior functional analysis. This is not a minor task given the large number of variants and genes that may occur in the inverted regions (e.g. the ruff supergene is expected to contain up to 125 different genes). Pedigree-free methods can also be applied to detect inversions (e.g. via comparisons of several assemblies [70]). Similarly, pleiotropic (or epistatic) effects can also be detected from genomic data by specifically modelling these effects during data analysis (e.g. [71]). Within the same studies, other factors leading to spurious genetic correlations (assortative mating, migration, admixture) between colour and other traits can be discarded as they are expected to result in divergence spreading across the genome.

In the past decades, numerous studies were done to investigate the condition dependency of several types of colorations [72]. The association between condition and coloration and its evolvability is expected to be based, at least in some evolutionary scenarios, on a genetic correlation between body condition and coloration [73]. In these cases, genetic variation in condition is expected to determine genetic variation on coloration, and the molecular basis of condition-dependent expression of coloration will not be expected on genes that are ultimately linked with the development of coloration

(e.g. pigment synthesis), unless they also determine condition. The main difficulties to unravel the condition dependency of certain colour traits could come from the expected highly polygenic basis of these traits [74] and the multifaceted nature of condition. A potential approach would be to 'divide' condition into the different aspects that may define it for a given species. For a trait based on carotenoids for example, we could expect that genetic variation at the level of foraging efficiency, absorption efficiency, storage allocation and immunological response defines, among others, the aspects of condition that actually determine the development of a given trait. Performance of individuals at each of these levels and its effect on coloration can be measured, and genome scans of individuals with the most differentiated levels of performance (and colour) can be conducted to search for candidate variants. The analysis of changes in gene expression with condition will also be of great utility to pinpoint the association between condition and colour regulation, and the different genetic actors that might be at play when different aspects of coloration are manipulated.

5. Delving into evolutionary mechanisms

Knowledge on the genetic architecture will be of major importance to clarify how animal colour traits evolve. Predicting phenotypic evolution in the wild has turned out to be a highly challenging task given the action of selection on other genetically correlated traits, physical linkage, and epistatic and dominant interactions between and within loci, respectively. Genomic tools can provide useful insights on these determinants (e.g. [15], figure 1*a*). They can help us to understand how colour traits respond to selection, how variation is maintained and the general applicability of the phenotypic gambit [13]. Predicting phenotypic evolution is also the ultimate goal of convergence and parallel evolution studies, which have ideal model studies in mimic butterflies [32], adaptive radiations of cichlid fishes [75] and adaptive melanism in rodents, lizards and birds [76]. Understanding the molecular means through which organisms evolve similar colour traits is possible with genomic tools without the biases associated with candidate gene approaches [5]. In this sense, it will also be important to unravel potentially constraining interactions between genes (i.e. epistatic and pleiotropic gene effects) to test whether molecular adaptation proceeds through the path of less genetic resistance, or whether selection can overcome such contingencies.

Specific gene effects or genetic architectures are also at the core of several evolutionary theories. Rowe & Houle [18] proposed that the capture of genetic variance by condition-dependent traits would help to solve how genetic variation for sexually selected traits is maintained, i.e. the so-called lek paradox. This model assumes a large genetic variance for condition, owing to many small-effect QTLs associated to condition, and consequently, a high mutation rate for condition [77]. These assumptions can be challenged at least in part by unravelling the genetic architecture of condition-dependent colour traits. Similarly, Hamilton & Zuk [78] proposed that individual coloration can signal an individual's genetic capacity to resist parasites, and that host-parasite co-adaptation cycles could maintain genetic variation in male traits and, consequently, female choice. Such dynamics can be tracked down knowing which causal variants underlie colour differences between individuals [79]. These are only a few of the theoretical frameworks underlying the study of animal coloration; other scenarios of 'good genes' affecting offspring viability have been proposed to drive indirect benefits of mate choice and its evolution [80], and the incomplete evolution of sexual dimorphism is predicted to result from loci that have opposing fitness effects on males and females (intra- and inter-locus sexual conflict; [81]). We believe that genomic studies can be designed with the ultimate goal of testing the genetic assumptions of these and other theories, which we predict will increase our understanding of the evolution of animal coloration.

6. Final remarks: validation and novelty

We have highlighted different frameworks where using high-throughput sequencing data could aid in advancing the study of animal coloration. It is important to equally highlight that there are plenty of scenarios where the application of genomic tools is not justified or can lack power (e.g. to detect lowly expressed genes or variants in genomic regions that are difficult to sequence). Previous approaches, such as the candidate gene approach, will be more useful, for instance, to address specific hypotheses on the role of certain genes [82]. Similarly, it is important to highlight that studies using genomic or transcriptomic-wide approaches will be rarely definitive, and that their utility largely relies on follow-up studies validating the causality of the discovered candidate variants or transcripts. Validation studies are essential for two different reasons. On one hand, high-throughput sequencing data is highly prone to false-positive signals owing to the massive amount of data and multiple hypothesis testing [83], sequencing biases [84] and potential flaws in the experimental design [58]. Additionally, several variants or transcripts can be discovered owing to linkage or co-expression, while only few may have a true effect on coloration. On the other hand, if validation is not implemented, the potential of genomic and transcriptomic studies to unravel novel genetic pathways will be undermined by the lack of experimental evidence on the functional role of the new discoveries.

Validation is not a minor task given that gathering robust experimental support of the functional effects of a given variant or gene can be difficult in non-model species. Closely related model species can be used to validate the causal function. Santos *et al*. [16] used this approach to validate the role of the gene *fhl2b* in mediating egg-spot formation in the anal fins of haplochromin cichlids (figure 1*b*). Using whole-transcriptome shotgun sequencing, Santos *et al*. found *fhl2b* to be highly expressed in anal fins presenting egg-spots. A comparison of the genomic sequences of different cichlids species revealed the presence of a transposable element in the regulatory region upstream the gene *fhl2b* in species displaying egg spots. They then generated transgenic lines of the closest model species, the zebrafish, which confirmed that the insertion of the transposon upregulates the expression of *fhl2b*. Other manipulative approaches like CRISPR/Cas9 [24], or the transfection of cell cultures [85], have also been successfully applied to validate the role of candidate variants on coloration.

Although desirable, manipulative experiments may not be feasible in certain species, or may not be conclusive (some variants may have a true effect but only on the genetic or physiological background where they were discovered). Thus, gathering multiple lines of evidence on the role of a gene or a

gene variant will be more informative and should become standard practice. Multiple lines of evidence can be obtained in most study systems by: (i) following the transmission of the candidate variants and coloration within a pedigree [86], (ii) confirming predicted effects on expression (e.g. measuring gene expression during colour development or between body parts expressing different colours [22,23]), (iii) replicating the study in independent samples or in related species exhibiting similar colour traits [16], or (iv) by assessing the capacity of the observed variants in predicting colour variation in a different sample [53]. Given the importance and difficulty of validating findings obtained with high-throughput sequencing data, we encourage discussion and consensus on this topic as recently done for association studies in humans [87]. Without the ultimate goal of validation, and further research on the function of newly as well as previously discovered genes, genome-wide tools will be largely misleading and of little help in understanding the wide and complex diversity of animal colour traits.

Authors' contributions. A.R. conceived the study, and L.M.S.-J. and A.R. wrote the manuscript.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the Swiss National Science Foundation (31003A-120517 to A.R.).

Acknowledgements. We dedicate this study to the memory of Dr Elaina M. Tuttle. We thank Vera Uva, Devi Stuart-Fox and two anonymous reviewers for helpful suggestions on previous versions of the manuscript.

References

1. Hoekstra HE. 2006 Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity* **97**, 222–234. (doi:10.1038/sj.hdy.6800861)
2. Kettlewell HBD. 1955 Selection experiments on industrial malanism in the *Lepidoptera*. *Heredity* **9**, 323–342. (doi:10.1038/hdy.1955.36)
3. Endler JA. 1983 Natural and sexual selection on color patterns in poeciliid fishes. *Environ. Biol. Fishes* **9**, 173–190. (doi:10.1007/bf00690861)
4. Singh AP, Nüsslein-Volhard C. 2015 Zebrafish stripes as a model for vertebrate colour pattern formation. *Curr. Biol.* **25**, R81–R92. (doi:10.1016/j.cub.2014.11.013)
5. Protas ME, Patel NH. 2008 Evolution of coloration patterns. *Annu. Rev. Cell Dev. Biol.* **24**, 425–446. (doi:10.1146/annurev.cellbio.24.110707.175302)
6. Caro T. 2005 The adaptive significance of coloration in mammals. *Bioscience* **55**, 125–136. (doi:10.1641/0006-3568(2005)055[0125:TASOCI]2.0.CO;2)
7. Roulin A. 2004 The evolution, maintenance and adaptive function of genetic colour polymorphism in birds. *Biol. Rev.* **79**, 815–848. (doi:10.1017/s1464793104006487)
8. Kronforst MR *et al.* 2012 Unraveling the thread of nature's tapestry: the genetics of diversity and convergence in animal pigmentation. *Pigment Cell Melanoma Res.* **25**, 411–433. (doi:10.1111/j.1755-148X.2012.01014.x)
9. Mundy NI, Badcock NS, Hart T, Scribner K, Janssen K, Nadeau NJ. 2004 Conserved genetic basis of a quantitative plumage trait involved in mate choice. *Science* **303**, 1870–1873. (doi:10.1126/science.1093834)
10. Hoekstra HE, Nachman MW. 2003 Different genes underlie adaptive melanism in different populations of rock pocket mice. *Mol. Ecol.* **12**, 1185–1194. (doi:10.1046/j.1365-294X.2003.01788.x)
11. Maynard-Smith J. 1978 Optimization theory in evolution. *Annu. Rev. Ecol. Syst.* **9**, 31–56. (doi:10.1016/0016-0032(73)90062-8)
12. Grafen A. 1990 Biological signals as handicaps. *J. Theor. Biol.* **144**, 517–546. (doi:10.1016/S0022-5193(05)80088-8)
13. Hadfield JD, Nutall A, Osorio D, Owens IPF. 2007 Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *J. Evol. Biol.* **20**, 549–557. (doi:10.1111/j.1420-9101.2006.01262.x)
14. da Fonseca RR *et al.* 2016 Next-generation biology: sequencing and data analysis approaches for non-model organisms. *Mar. Genomics* **30**, 3–13. (doi:10.1016/j.margen.2016.04.012)
15. Comeault AA, Ferreira C, Dennis S, Soria-Carrasco V, Nosil P. 2016 Color phenotypes are under similar genetic control in two distantly related species of Timema stick insect. *Evolution* **70**, 1283–1296. (doi:10.1017/CBO9781107415324.004)
16. Santos ME, Braasch I, Boileau N, Meyer BS, Sauteur L, Böhne A, Belting H-G, Affolter M, Salzburger W. 2014 The evolution of cichlid fish egg-spots is linked with a *cis*-regulatory change. *Nat. Commun.* **5**, 5149. (doi:10.1038/ncomms6149)
17. Küpper C *et al.* 2015 A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.* **48**, 79–83. (doi:10.1038/ng.3443)
18. Poelstra JW, Vijay N, Hoeppner MP, Wolf JBW. 2015 Transcriptomics of colour patterning and coloration shifts in crows. *Mol. Ecol.* **24**, 4617–4628. (doi:10.1111/mec.13353)
19. Lopes RJ *et al.* 2016 Genetic basis for red coloration in birds. *Curr. Biol.* **26**, 1427–1434. (doi:10.1016/j.cub.2016.03.076)
20. Mundy NI *et al.* 2016 Red carotenoid coloration in the zebra finch is controlled by a cytochrome P450 gene cluster. *Curr. Biol.* **26**, 1435–1440. (doi:10.1016/j.cub.2016.04.047)
21. Kardos M, Husby A, Mcfarlane SE, Qvarnström A, Ellegren H. 2016 Whole-genome resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative trait loci in natural populations. *Mol. Ecol. Resour.* **16**, 727–741. (doi:10.1111/1755-0998.12498)
22. Nadeau NJ *et al.* 2016 The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature* **534**, 106–110. (doi:10.1038/nature17961)
23. Gallant JR *et al.* 2014 Ancient homology underlies adaptive mimetic diversity across butterflies. *Nat. Commun.* **5**, 4817. (doi:10.1038/ncomms5817)
24. Li X *et al.* 2015 Outbred genome sequencing and CRISPR/Cas9 gene editing in butterflies. *Nat. Commun.* **6**, 8212. (doi:10.1038/ncomms9212)
25. Dembeck LM, Huang W, Magwire MM, Lawrence F, Lyman, RF, Mackay TFC. 2015 Genetic architecture of abdominal pigmentation in *Drosophila melanogaster*. *PLOS Genet.* **11**, e1005163. (doi:10.1371/journal.pgen.1005163)
26. Bradbury JW, Vehrencamp SL. 2011 Light and visual signal production. In *Principles of animal communication*, pp. 113–152. Sunderland, MA: Sinauer Associates, Inc.
27. Takeuchi K, Satou Y, Yamamoto H, Satoh N. 2005 A genome-wide survey of genes for enzymes involved in pigment synthesis in an ascidian, *Ciona intestinalis*. *Zoolog. Sci.* **22**, 723–734. (doi:10.2108/zsj.22.723)
28. Lin SJ *et al.* 2013 Topology of feather melanocyte progenitor niche allows complex pigment patterns to emerge. *Science* **340**, 1442–1445. (doi:10.1126/science.1230374)
29. True JR. 2003 Insect melanism: the molecules matter. *Trends Ecol. Evol.* **18**, 640–647. (doi:10.1016/j.tree.2003.09.006)
30. Wittkopp PJ, Carroll SB, Kopp A. 2003 Evolution in black and white: genetic control of pigment patterns in *Drosophila*. *Trends Genet.* **19**, 495–504. (doi:10.1016/S0168-9525(03)00194-X)
31. Bourgeois YXC, Bertrand JAM, Delahaie B, Cornuault J, Duval T, Milá B, Thébaud C. 2016 Candidate gene analysis suggests untapped genetic complexity in melanin-based pigmentation in birds. *J. Hered.* **107**, 327–335. (doi:10.1093/jhered/esw017)
32. Kronforst MR, Papa R. 2015 The functional basis of wing patterning in *Heliconius* butterflies: the molecules behind mimicry. *Genetics* **200**, 1–19. (doi:10.1534/genetics.114.172387)
33. Roberts RB, Ser JR, Kocher TD. 2009 Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *Science* **326**, 998–1001. (doi:10.1126/science.1174705)
34. Crookes WJ, Ding LL, Huang QL, Kimbell J, Horwitz J, McFall-Ngai MJ. 2004 Reflectins: the unusual proteins of squid reflective tissues. *Science* **303**, 235–238. (doi:10.1126/science.1091288)

35. Kemp DJ, Herberstein ME, Grether GF. 2011 Unraveling the true complexity of costly color signaling. *Behav. Ecol*. **23**, 233–236. (doi:10.1093/beheco/arr153)
36. San-Jose LM, Granado-Lorencio F, Sinervo B, Fitze PS. 2013 Iridophores and not carotenoids account for chromatic variation of carotenoid-based coloration in common lizards (*Lacerta vivipara*). *Am. Nat*. **181**, 396–409. (doi:10.1086/669159)
37. Saenko SV, Teyssier J, van der Marel D, Milinkovitch MC. 2013 Precise colocalization of interacting structural and pigmentary elements generates extensive color pattern variation in Phelsuma lizards. *BMC Biol*. **11**, 105. (doi:10.1186/1741-7007-11-105)
38. Manceau M, Domingues VS, Mallarino R, Hoekstra HE. 2011 The developmental role of Agouti in color pattern evolution. *Science* **331**, 1062–1065. (doi:10.1126/science.1200684)
39. Shorter KR, Crossland JP, Webb D, Szalai G, Felder MR, Vrana PB. 2012 Peromyscus as a mammalian epigenetic model. *Genet. Res. Int*. **2012**, 179159. (doi:10.1155/2012/179159)
40. Faulk C, Dolinoy DC. 2011 Timing is everything: the when and how of environmentally induced changes in the epigenome of animals. *Epigenetics* **6**, 791–797. (doi:10.4161/epi.6.7.16209)
41. Santos ESA, Scheck D, Nakagawa S. 2011 Dominance and plumage traits: meta-analysis and metaregression analysis. *Anim. Behav*. **82**, 3–19. (doi:10.1016/j.anbehav.2011.03.022)
42. Wu AR *et al.* 2014 Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* **11**, 41–46. (doi:10.1038/nmeth.2694)
43. Michels KB *et al.* 2013 Recommendations for the design and analysis of epigenome-wide association studies. *Nat. Methods* **10**, 949–955. (doi:10.1038/nmeth.2632)
44. Ekblom R, Wolf JBW. 2014 A field guide to whole-genome sequencing, assembly and annotation. *Evol. Appl*. **7**, 1026–1042. (doi:10.1111/eva.12178)
45. Vijay N, Poelstra JW, Künstner A, Wolf JBW. 2013 Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive in silico assessment of RNA-seq experiments. *Mol. Ecol*. **22**, 620–634. (doi:10.1111/mec.12014)
46. Hoban S *et al.* 2016 Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *Am. Nat*. **188**, 379–397. (doi:10.1086/688018)
47. Wang S, Gribskov M. 2017 Comprehensive evaluation of de novo transcriptome assembly programs and their effects on differential gene expression analysis. *Bioinformatics* **33**, 327–333. (doi:10.1093/bioinformatics/btw625)
48. Card DC *et al.* 2014 Two low coverage bird genomes and a comparison of reference-guided versus de novo genome assemblies. *PLoS ONE* **9**, e106649. (doi:10.1371/journal.pone.0106649)
49. Roulin A. 2016 Condition-dependence, pleiotropy and the handicap principle of sexual selection in melanin-based colouration. *Biol. Rev*. **91**, 328–348. (doi:10.1111/brv.12171)
50. Svensson PA, Wong BBM. 2011 Carotenoid-based signals in behavioural ecology: a review. *Behaviour* **148**, 131–189. (doi:10.1163/000579510X548673)
51. Mackay TFC, Stone EA, Ayroles JF. 2009 The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet*. **10**, 565–577. (doi:10.1038/nrg2612)
52. Barton NH, Keightley PD. 2002 Understanding quantitative genetic variation. *Nat. Rev. Genet*. **3**, 11–21. (doi:10.1038/nrg700)
53. Liu F, van Duijn K, Vingerling JR, Hofman A, Uitterlinden AG, Janssens ACJW, Kayser M. 2009 Eye color and the prediction of complex phenotypes from genotypes. *Curr. Biol*. **19**, 192–193. (doi:10.1016/j.cub.2009.01.027)
54. Slate J. 2005 Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Mol. Ecol*. **14**, 363–379. (doi:10.1111/j.1365-294X.2004.02378.x)
55. Yang J *et al.* 2010 Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet*. **42**, 565–569. (doi:10.1038/ng.608)
56. Endler L, Betancourt AJ, Nolte V, Schlötterer C. 2016 Reconciling differences in pool-GWAS between populations: a case study of female abdominal pigmentation in *Drosophila melanogaster*. *Genetics* **202**, 843–855. (doi:10.1534/genetics.115.183376)
57. Richards PM, Liu MM, Lowe N, Davey JW, Blaxter ML, Davison A. 2013 RAD-Seq derived markers flank the shell colour and banding loci of the *Cepaea nemoralis* supergene. *Mol. Ecol*. **22**, 3077–3089. (doi:10.1111/mec.12262)
58. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, Geman D, Baggerly K, Irizarry RA. 2010 Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet*. **11**, 733–739. (doi:10.1038/nrg2825)
59. Amos CI. 2007 Successful design and conduct of genome-wide association studies. *Hum. Mol. Genet*. **16**, 220–225. (doi:10.1093/hmg/ddm161)
60. Slate J, Pemberton JM, Visscher PM. 1999 Power to detect QTL in a free-living polygynous population. *Heredity* **83**, 327–336. (doi:10.1038/sj.hdy.6885830)
61. Yong L, Peichel CL, McKinnon JS. 2015 Genetic architecture of conspicuous red ornaments in female threespine stickleback. *G3* **6**, 579–588. (doi:10.1534/g3.115.024505)
62. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES. 2009 Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* **21**, 2194–2202. (doi:10.1105/tpc.109.068437)
63. McKinnon JS, Pierotti MER. 2010 Colour polymorphism and correlated characters: genetic mechanisms and evolution. *Mol. Ecol*. **19**, 5101–5125. (doi:10.1111/j.1365-294X.2010.04846.x)
64. Sinervo B, Svensson E. 2002 Correlational selection and the evolution of genomic architecture. *Heredity* **89**, 329–338. (doi:10.1038/sj.hdy.6800148)
65. Ducrest A-L, Keller L, Roulin A. 2008 Pleiotropy in the melanocortin system, coloration and behavioural syndromes. *Trends Ecol. Evol*. **23**, 502–510. (doi:10.1016/j.tree.2008.06.001)
66. Wilkins AS, Wrangham RW, Tecumseh Fitch W. 2014 The 'domestication syndrome' in mammals: a unified explanation based on neural crest cell behavior and genetics. *Genetics* **197**, 795–808. (doi:10.1534/genetics.114.165423)
67. Wittkopp PJ, Beldade P. 2009 Development and evolution of insect pigmentation: genetic mechanisms and the potential consequences of pleiotropy. *Semin. Cell Dev. Biol*. **20**, 65–71. (doi:10.1016/j.semcdb.2008.10.002)
68. Nishikawa H *et al.* 2015 A genetic mechanism for female-limited Batesian mimicry in *Papilio* butterfly. *Nat. Genet*. **47**, 405–409. (doi:10.1038/ng.3241)
69. Sih A, Bell A, Johnson J, Ziemba R. 2004 Behavioural syndromes: an integrative overview. *Q. Rev. Biol*. **79**, 241–277. (doi:10.1086/422893)
70. Tattini L, D'Aurizio R, Magi A. 2015 Detection of genomic structural variants from next-generation sequencing data. *Front. Bioeng. Biotechnol*. **3**, 92. (doi:10.3389/fbioe.2015.00092)
71. Korte A, Vilhjálmsson BJ, Segura V, Platt A, Long Q, Nordborg M. 2012 A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nat. Genet*. **44**, 1066–1071. (doi:10.1038/ng.2376)
72. Hill GE. 2011 Condition-dependent traits as signals of the functionality of vital cellular processes. *Ecol. Lett*. **14**, 625–634. (doi:10.1111/j.1461-0248.2011.01622.x)
73. Wilkinson GS, Taper M. 1999 Evolution of genetic variation for condition-dependent traits in stalk-eyed flies. *Proc. R. Soc. Lond. B* **266**, 1685–1690. (doi:10.1098/rspb.1999.0832)
74. Rowe L, Houle D. 1996 The lek paradox and the capture of genetic variance by condition dependent traits. *Proc. R. Soc. Lond. B* **263**, 1415–1421. (doi:10.1098/rspb.1996.0207)
75. Elmer KR, Meyer A. 2011 Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol. Evol*. **26**, 298–306. (doi:10.1016/j.tree.2011.02.008)
76. Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE. 2010 Convergence in pigmentation at multiple levels: mutations, genes and function. *Phil. Trans. R. Soc. B* **365**, 2439–2450. (doi:10.1098/rstb.2010.0104)
77. Tomkins JL, Radwan J, Kotiaho JS, Tregenza T. 2004 Genic capture and resolving the lek paradox. *Trends Ecol. Evol*. **19**, 323–328. (doi:10.1016/j.tree.2004.03.029)
78. Hamilton WD, Zuk M. 1982 Heritable true fitness and bright birds: a role for parasites? *Science* **218**, 384–387. (doi:10.1126/science.7123238)
79. Balenger SL, Zuk M. 2014 Testing the Hamilton-Zuk hypothesis: past, present, and future. *Integr. Comp. Biol*. **54**, 601–613. (doi:10.1093/icb/icu059)
80. Kokko H, Jennions MD, Brooks R. 2006 Testing and unifying models of sexual selection. *Annu. Rev. Ecol.*

Evol. Syst. **37**, 43–66. (doi:10.2307/annurev.ecolsys.37.091305.30000004)

81. van Doorn GS. 2009 Intralocus sexual conflict. *Ann. N. Y. Acad. Sci*. **1168**, 52–71. (doi:10.1111/j.1749-6632.2009.04573.x)
82. San-Jose LM, Ducrest AL, Ducret V, Simon C, Richter H, Wakamatsu K, Roulin A. 2017 MC1R variants affect the expression of melanocortin and melanogenic genes and the association between melanocortin genes and coloration. *Mol. Ecol*. **26**, 259–276. (doi:10.1111/mec.13861)
83. Lambert CG, Black LJ. 2012 Learning from our GWAS mistakes: from experimental design to scientific method. *Biostatistics* **13**, 195–203. (doi:10.1093/biostatistics/kxr055)
84. Robasky K, Lewis NE, Church GM. 2014 The role of replicates for error mitigation in next-generation sequencing. *Nat. Rev. Genet*. **15**, 56–62. (doi:10.1038/nrg3655)
85. Rosenblum EB, Römpler H, Schöneberg T, Hoekstra HE. 2010 Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proc. Natl Acad. Sci. USA* **107**, 2113–2117. (doi:10.1073/pnas.0911042107)
86. Gratten J, Wilson AJ, McRae AF, Beraldi D, Visscher PM, Pemberton JM, Slate J. 2008 A localized negative genetic correlation constrains microevolution of coat color in wild sheep. *Science* **319**, 318–320. (doi:10.1126/science.1151182)
87. MacArthur DG *et al.* 2014 Guidelines for investigating causality of sequence variants in human disease. *Nature* **508**, 469–476. (doi:10.1038/nature13127)