CellPress
OPEN ACCESS

Trends in
Ecology & Evolution

Review

Functional genomic tools for emerging model species

Erik Gudmunds,[1,*] Christopher W. Wheat,[2] Abderrahman Khila,[1,3] and Arild Husby[1,*]

Most studies in the field of ecology and evolution aiming to connect genotype to phenotype rarely validate identified loci using functional tools. Recent developments in RNA interference (RNAi) and clustered regularly interspaced palindromic repeats (CRISPR)-Cas genome editing have dramatically increased the feasibility of functional validation. However, these methods come with specific challenges when applied to emerging model organisms, including limited spatial control of gene silencing, low knock-in efficiencies, and low throughput of functional validation. Moreover, many functional studies to date do not recapitulate ecologically relevant variation, and this limits their scope for deeper insights into evolutionary processes. We therefore argue that increased use of gene editing by allelic replacement through homology-directed repair (HDR) would greatly benefit the field of ecology and evolution.

Highlights

Understanding the molecular mechanisms underlying phenotypic evolution is a central goal in evolutionary biology.

Unfortunately, causal associations between candidate loci and the phenotype are rarely established because the majority of associations are never validated with functional tools.

We describe the latest developments in the use of functional tools, particularly RNAi and clustered regularly interspaced palindromic repeats (CRISPR)-Cas, to establish causal links between genotype and phenotype, and discuss some methodological challenges when applying these tools that are unique to emerging model organisms.

Specifically, we highlight a significant issue in that current functional studies in ecology and evolutionary biology to date rarely recapitulate ecologically relevant variation, and we suggest that gene editing in emerging model species should make more use of the HDR pathway to obtain relevant functional insights and thereby a deeper understanding of the genetic basis of traits.

Identifying genomic variation in evolutionary dynamics

A fundamental aim in ecology and evolution is to identify the genomic variation that underlies micro- and macroevolutionary dynamics [1]. In the present post-genomic era, diverse tools and datasets are routinely used for this purpose in **emerging model organisms** (see Glossary), ranging from genome-wide population data and various phenotype association approaches to differential expression analysis and chromatin accessibility assays [2,3]. Although such methods are advancing our understanding of genotype–phenotype relationships for traits involved in adaptation and speciation [3,4], they also have clear limitations, the most obvious being that they lack the ability to establish rigorous causal inferences between phenotypes and **candidate genes** or loci. However, molecular tools for manipulating gene function, such as RNAi and clustered regularly interspaced palindromic repeats (CRISPR)–CRISPR-associated protein (Cas), in recent years have also made it possible to establish a causal link between candidate loci and phenotype in emerging model organisms. In this review we address the growing need for such causal inferences, and we discuss some of the practical aspects of **functional validation** in emerging model organisms and the future role of RNAi and CRISPR-Cas technologies in ecology and evolution, with a strong emphasis on the challenge and importance of establishing ecological relevance in functional validation.

The expanded reach of functional genomics

Rather than investing in functional validation, the accessibility of **functional genomics** hasinstead predominantly resulted in an increase in the sample sizes of individuals, populations, and species used in the quest to identify loci involved in adaptation and speciation. The result is an ever-increasing number of studies that, after identifying candidate loci (e.g., through molecular tests of selection, allele-frequency patterns, association mapping, linkage analysis, and transcriptomic studies), discuss at length the potential evolutionary implications. These methods all have inherent challenges (reviewed in [5–8]), but common to them is the limitation that they do not establish causal associations between genotype and phenotype. Drawing rigorous conclusions

[1]Evolutionary Biology, Department of Ecology and Genetics, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden
[2]Department of Zoology, Stockholm University, Svante Arrhenius väg 18B, 10691, Stockholm, Sweden
[3]Institut de Génomique Fonctionelle de Lyon, Université de Lyon, Université Claude Bernard Lyon, Centre, National de la Recherche Scientifique (CNRS) Unité Mixte de Recherche 5242, École Normale Supérieure de Lyon, 46 allée d'Italie, 69364, Lyon CEDEX 07, France

*Correspondence:
erik.gudmunds@ebc.uu.se (E. Gudmunds) and
arild.husby@ebc.uu.se (A. Husby).

from this body of evidence is therefore challenging, and interpretations of such insights should be considered carefully when attempting to understand the genotype–phenotype relationship.

Given these circumstances, we raise two aspects worth considering when investigating the functional genetic basis of traits in ecology and evolution. First, although we acknowledge that the implementation of functional genetic methods is not possible or even desirable in some instances, we wish to stress that, in the absence of proper data allowing causal inferences, the field would greatly benefit from avoiding the use of causal language. Imprecise language degrades the importance of causal insights, impedes our understanding of how the natural world works, and bloats the field with false positive results [9]. Second, relying upon existing annotations without validation perpetuates bias. Consider a common hypothetical scenario, wherein candidate loci are identified using the intersection of gene expression data and SNP association mapping for a trait, where variants associated with the phenotype are in or near the coding sequences of several differentially expressed genes. The resulting candidate gene list is assessed, and hypotheses formed based on gene or sequence element annotations, which in turn are overwhelmingly based on genetic and molecular studies in the closest **model organism** (e.g., *Drosophila melanogaster* for insects, the mouse *Mus musculus* for mammals, the zebrafish *Danio rerio* for fish, *Saccharomyces cerevisiae* for yeast, *Arabidopsis thaliana* for plants, and so on). We note the duality of this approach: although this seems a sound way forward because many homologous genes share functions over deep evolutionary time, and this approach is necessary because no other such insights are possible, it also comes with a risk of erroneous inferences. Model organisms not only have many important genes of unknown function [10] but functions also certainly change over time for orthologous genes [11]. Thus, relying solely on functional annotation from other organisms can restrict the ability to detect lineage-specific protein functions or genes, **moonlighting proteins**, or cases of **gene co-option** that have evolved to shape the phenotype of interest, especially because subsequent changes to functional annotations can even alter the biological conclusions [12]. Equally problematic, relying upon functional annotations for selecting candidate genes generates ascertainment bias because genes without assigned functions are almost always avoided, even though species-specific genes can clearly be important [11]. The magnitude of this problem is huge and the ramifications vast – even in *Homo sapiens* <10% of genes have been the focus of >90% of published papers [13].

We highlight this annotation bias, as well as the aforementioned issues, to demonstrate not only the importance but also the empowerment of being able to functionally test candidate genes irrespective of whether they have or lack annotations. A recent study by van der Burg *et al.* [3] exemplifies all of these issues with elegance, wherein the authors investigated the genetic assimilation of plasticity in butterfly wing coloration, arriving at 13 candidate loci after intersecting genome-wide association study (GWAS) and RNA-seq time-series results. They then used CRISPR-Cas to manipulate four of these candidate genes. **Knockout (KO)** of their first choice, a locus annotated to butterfly wing-pattern function (*cortex*), predictably affected wing color. Of the two other annotated genes, one had no phenotypic effect (*Dscam3*) whereas the other did (*trehalase*). Finally, KO of the then uncharacterized and unannotated gene demonstrated it to be causally involved in wing coloration plasticity, and it was named *herfst*. Had the authors chosen to only KO *cortex*, two novel loci would have been missed, one of which (the central metabolic gene *trehalase)* has an unpredicted role in wing patterning. Had they not conducted functional validation, the role of the neuronal developmental gene *Dscam3* in wing patterning might have generated a line of inquiry that in fact lacks causal support, and the novel role of *trehalase* might possibly have been discounted. Finally, their discovery of *herfst* not only provides important knowledge for future studies but also allows reanalysis of previous work. With more validation of

Glossary

Candidate gene: a gene or locus which has been associated with a phenotype of interest, for example through molecular test of selection, genome-wide association studies (GWAS), quantitative trait locus (QTL) mapping, or RNA-seq, and for which a hypothesis of a role in the phenotype of interest is formulated.
Ecologically relevant variation: when gene editing recreates phenotypic or genetic variation that is already present in the population as opposed to introducing novel variation.
Emerging model organism: an organism that lacks a substantial level of resources such as mutant libraries, large-scale screens for gene function, in-depth description of development and physiology, data-sharing communities (such as FlyBase, WormBase, and The *Arabidopsis* Information Resource, TAIR), and large specialized molecular toolboxes for genetic manipulation. These are also called non-model organisms or ecological model organisms. The recent access to genomic resources and the possibility of functional work and rearing in the laboratory are removing the divide between model organisms and 'non-model organisms', and thus 'emerging model organisms' is preferable to 'non-model organism'.
Functional genomics: the area of research that aims to understand how genes and intergenic regions contribute to biological processes.
Functional validation: using functional genetic methods that have the ability to demonstrate causality to test whether a candidate gene has a role in shaping a phenotype.
Gene editing: editing a locus from one allele to another (does not need to be a naturally occurring allele).
Genetic compensation: a form of canalization of the phenotype which can mask disruptive mutations. It can be detected by comparing KO and knockdown effects of the same gene. Loss of function of one gene is compensated by altered expression of other genes.
Genetic co-option: the employment of conserved gene functions or pathways in a new process, for example the formation of new traits.
Homology-directed repair (HDR): a ubiquitous DNA repair system which repairs double-strand breaks (DSBs)

components of studies in ecology and evolution, these types of functional insights will synergistically improve future association studies across a broader taxonomical range for phenotypes of interest to our field.

Functional validation of candidate loci in emerging model organisms

Establishing functional tools for an emerging model system is by necessity a stepwise process that requires determination and effort. Although some species will never be amenable to functional manipulation, validation in others might be accomplished on the scale of months, which has led to an impressive proliferation of insightful studies using gene **knockdown** or KO techniques to explore the genetic basis of traits from candidates that were or had previously been identified from association-based approaches [14–20]. These studies are taxonomically diverse, ranging from confirming the role of loci in adaptive wing coloration in butterflies [15,16,20] and albinism in cave fish [14] to describing the role of lineage-specific genes in adaptation to new habitats in water striders [17]. These findings also highlight one of the major benefits of bringing functional tools to the realm of emerging model organisms – the possibility to explore the molecular mechanisms behind evolutionary processes that are documented in natural populations but that are difficult to study in model organisms. These include phenomena such as the genetic basis of migration [21], trait exaggeration [22], evolutionary innovations [23], and polyphenisms [24], to name but a few.

The tools – RNAi and its use in emerging model organisms

RNAi is a molecular technique that is used to artificially downregulate the expression of a gene of interest to assess its functional importance at the phenotypic level. The downregulation of the gene occurs temporarily and is usually referred to as a gene knockdown because it reduces rather than eliminates protein production, as opposed to a gene KO. The use of RNAi largely relies on proteins of the endogenous small RNA pathways (e.g., microRNA, small interfering RNA, Piwi-interacting RNA) that are present in almost all eukaryotic lineages [25]. Although the functions and dynamics of endogenous small RNA pathways across taxa are highly divergent and complex [26], exogenously induced RNAi is usually considered to follow a simplified conserved mechanism (Box 1) which makes RNAi a possibility for most eukaryotic emerging model organisms (although there are notable exceptions; e.g., Lepidoptera [27]).

Box 1. The nature of the RNAi pathway and how it enables sequence-specific knockdowns

The RNAi pathway was first discovered in plants and *Caenorhabditis elegans* [78,79], where researchers observed a sequence-specific reduction of gene expression when they introduced double-stranded RNA (dsRNA) or antisense RNA (asRNA) into cells or whole organisms. Research in different model systems has since elucidated the molecular mechanisms of RNAi [80]. The core pathway is initiated by cleavage of dsRNA by the endonuclease Dicer into smaller 20–25 nt fragments [81]. From these small RNA fragments, RNA strands associate with proteins of the Argonaut family and, together with additional factors, form the **RNA-induced silencing complex (RISC)** [82]. Based on the sequence of the small RNA, the RISC recognizes and binds to other RNA molecules and interferes with their expression by direct endonucleolytic cleavage, translational repression, or by directing the formation of heterochromatin at the corresponding genomic locus [45].

The use of RNAi is dependent on the ability of the organism to transport RNA into the cytosol where the mechanisms described above are initiated. In some groups the transport of dsRNA is highly efficient, whereas in other groups such as vertebrates, dsRNA transport across cell membranes does not occur easily. In such cases, dsRNA introduction can be aided by exogenous transport agents or methods including viral vectors and electroporation, or delivering the dsRNA with the use of vehicle molecules such as cationic lipids [45]. A common technique for dsRNA introduction into emerging model organisms is microinjection, where the dsRNA in solution is injected directly into the tissue. This method is dependent on whether the organism has an efficient endogenous system for dsRNA transport, and is usually referred to as systemic RNAi. The strategy of delivering dsRNA into the relevant emerging model organism can, to a large extent, be guided by the methods applied in the nearest model organism.

with the aid of the complementary DNA sequence as a repair template.
Knockdown: reduction of gene expression caused by artificially induced RNA degradation or inhibition of protein translation.
Knockout (KO): the generation of gene variants which are not functional, usually by disruption of the coding sequence.
Model organism: an organism that has wide laboratory usage, many genetic resources, many different and often specialized tools for functional research, and is amenable to propagation in laboratory settings.
Moonlighting proteins: proteins that have one main, usually basic, cellular function, and another, usually more specialized, function in an unrelated process. Different parts of the proteins can be responsible for the different functions.
Non-homologous end-joining (NHEJ): a DNA repair pathway that repairs DSBs in DNA by joining the two ends, is error-prone, and often results in mutations.
Protospacer adjacent motif (PAM): a short (usually 3 bp) sequence which is needed for endonucleolytic cleavage by CRISPR-associated proteins (Cas proteins).
RNA-induced silencing complex (RISC): recognizes and binds to other RNA molecules and interferes with their expression.
Single guide RNA (sgRNA): the RNA that guides Cas proteins to the site of cleavage; one part that is programmed by the user and another part binds to Cas protein. A chimera between the prokaryotic CRISPR (cr)RNA and *trans*-acting CRISPR (tracr)RNA.

Advantages and limitations of RNAi in emerging model organisms

Apart from the generally broad applicability of RNAi in diverse taxa, additional advantages include initiation at specific developmental stages and the potential for tissue localization, both of which can minimize negative pleiotropic effects. Further, RNAi can be tuned by controlling the double-stranded RNA (dsRNA) dose to levels where the gene expression is not completely ablated [28,29], thus enabling survival and phenotyping so as to infer function. For example, the function of an insulin receptor in insect wing polyphenism has been revealed by RNAi [30], whereas KO alleles cause embryonic lethality in the homozygous state [31]. Finally, RNAi has also been used in several emerging model organisms to study polyphenisms [32] which have rarely been studied in classical model organisms.

Although large-scale RNAi screens for gene function are normally associated with classical model organisms [33,34], screens of tens of genes are possible in emerging model organisms [35]. Efficient screens rely on overcoming the hurdle of dsRNA introduction into the organism, which in most cases is done through time-consuming microinjections, but targeting a modest set of genes (less than ~50) is achievable through microinjection techniques [17,35]. Some emerging model organisms are amenable to feeding RNAi [36,37], which enables time-efficient knockdowns of many genes [38]. In addition to feeding RNAi, topical application of dsRNA in insects [39] is an underutilized strategy for RNAi screens. One potential obstacle with RNAi is the efficiency with which dsRNA delivery and transport occurs within a treated individual because this varies across species [27] and even populations [40], generating variation that necessitates larger sample sizes.

Another drawback is that, in organisms lacking systemic RNAi, administration of dsRNA for efficient knockdown is difficult. One strategy in vertebrates is to resort to delivery via viral vectors [41], which has successfully been adopted in zebra finch to study the neurobiology of sex-specific song systems [42–44]. In addition to virus-mediated delivery, localized dsRNA injections can be used in vertebrates [45], which is a promising alternative when the tissue-specific expression pattern of the gene of interest is known or when the phenotype is manifested in clearly defined morphological structures. Finally, in organisms with long lifespans or long developmental times, the transient effect of RNAi can become a limitation and might require multiple or continuous administrations of dsRNA [46], although injection trauma itself in some cases can induce phenotypic changes [47]. In general, ontogenetic knowledge of the trait of interest is (always) useful and can be used to specifically guide the delivery of dsRNA to around the time when development of the trait occurs.

CRISPR-Cas – a versatile tool that potentiates a deeper understanding of phenotypes in emerging model species

The CRISPR-Cas toolbox (Box 2) is broad and has significantly increased the potential to perform functional validation in emerging model organisms [48]. Starting from the first description of **gene editing** with the CRISPR-Cas system [49], harnessing its full potential is a rapidly developing field. In addition to gene editing, researchers have been able to use CRISPR-Cas to specifically activate or repress gene expression [50,51], manipulate DNA methylation patterns [52], modify chromatin state [53], and visualize spatial patterns of chromatin conformation [54]. Although only a narrow range of these CRISPR tools currently work in emerging model organisms, all these levels of gene regulation play important roles in shaping phenotypes and may be important mechanisms of adaptation [55], making CRISPR-Cas poised to play an important role for functional validation of ecologically and evolutionarily relevant phenotypes in the future.

Box 2. Programmed genome editing with CRISPR-Cas

The ability of the CRISPR-Cas system to induce a targeted double-strand DNA break (DSB) has been well reviewed elsewhere [48,83,84], and we focus here upon what takes place after the DSB has been induced, usually 2–3 bp 5′ of the **protospacer adjacent motif (PAM)**; but this differs for different Cas endonucleases [85]). The DSB in the DNA can be repaired by one of two cell-autonomous DNA repair pathways: non-homologous end-joining (NHEJ) or homology-directed repair (HDR). The NHEJ pathway (Figure 1) identifies the DSB and joins the two DNA ends in an unpredictable error-prone fashion, which induces mutations in the form of insertions or deletions adjacent to the cut site. These resulting indels are relied upon for altering the reading frame when a DSB is targeted to the coding sequence (CDS) of a gene. However, when two flanking DSBs are induced by the use of two separate sgRNAs, deletion of entire regions can be accomplished. The HDR pathway by contrast uses homologous sequences as a template for the repair. The homologous sequences can either be endogenous, such as the homologous chromosome, or they can be exogenously introduced to the cell. In the latter case, gene editing occurs by including desired sequences within the HDR template (Figure 1), for example fluorescent genes, alternative alleles, and *cis*-regulatory elements (CREs), such that the original sequence is effectively replaced by the exogenous sequence. Depending on the CRISPR-Cas system, the material requirements are somewhat different, although the basics are very similar and are as easy to use as a reverse transcription kit, pushing deployment challenges to the realms of delivery, cut efficiency, and animal husbandry [86,87]. For some organisms, microinjection of reagents at developmentally relevant stages is greatly facilitated by soft eggs, external fertilization, and/or external development such as in some species of fish [88]. However, for species with hard eggs, such as many terrestrial arthropods, microinjection is difficult and can require substantial optimization [89].

The challenges of applying CRISPR-Cas in emerging model organisms are met with novel solutions

The starting point and end-goal of a gene-editing study dictates how to utilize the CRISPR-Cas toolbox. In principle, there are two distinct molecular categories of gene editing based on two different DNA repair pathways (Figure 1 and Box 2). Targeted mutagenesis through **non-homologous end-joining (NHEJ)** can be used to produce KOs for genes or regulatory regions. To exchange one allele for another one must instead rely on the **homology-directed repair (HDR)** pathway, which occurs less commonly in the cell and thus is less frequent than targeted mutagenesis via the NHEJ pathway. However, the HDR pathway enables allelic replacement experiments, something which is attractive for functional validation of quantitative trait loci (QTLs) [56] and the assessment of **ecologically relevant variation**, as compared to complete KO of gene function. Recent works have also successfully performed knock-ins by relying on homology-independent repair mechanisms in model organisms (*Drosophila* [57] and zebrafish [58]), and these may also be of use in emerging model organisms in the future.

Common to both types of gene editing is the design of **single guide RNAs (sgRNAs)** which can be accomplished with the aid of software that enables both the search for suitable target sequences (sgRNA design) in the sequence of interest and scanning of the rest of the genome for potential off-target sites [59]. Without a complete genome assembly, a powerful off-target control is to use several sgRNA constructs where each is designed to cut a different site in the locus of interest. When used independently, although any one sgRNA might induce off-target cleavage elsewhere in the genome, unique sgRNAs are unlikely to produce similar phenotypic effects unless they KO the function of the same gene. A final consideration is that combinations of sgRNAs can be injected. This is an excellent starting point to quickly determine if any of them have phenotypic effects because a given sgRNA may lack phenotypic effects for diverse reasons (e.g., secondary structures that prevent cuts, cuts do not ablate relevant gene function, etc). Using multiple sgRNAs is also important when targeting regulatory regions because a single sgRNA is often not sufficient to disrupt function.

Similar to the application of RNAi in emerging model organisms, the delivery of reagents into the desired cells is a significant challenge in using CRISPR-Cas (Box 3).

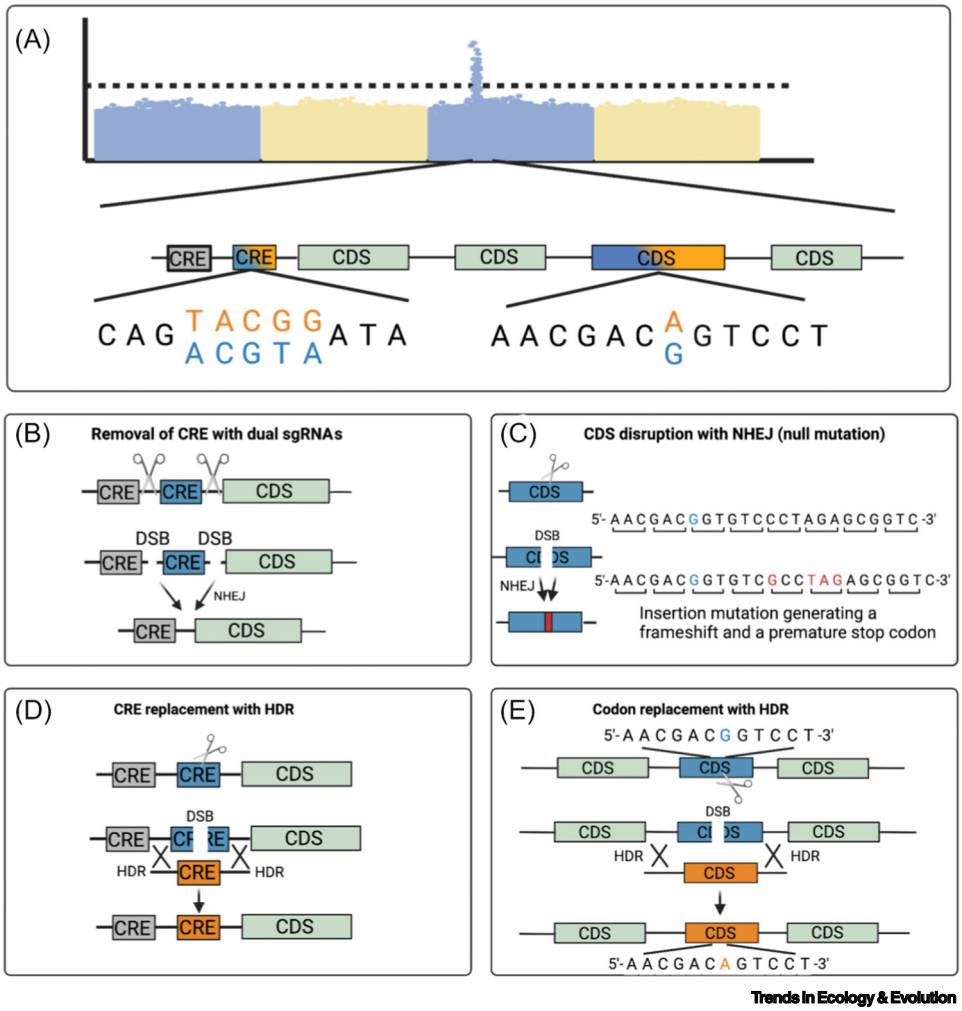

Figure 1. Genomic regions associated with phenotypes of interest are causal hypotheses that can be tested using CRISPR/Cas9 gene manipulations, some of which do and do not allow for ecologically relevant assessments. (A) Genetic variation in a cis-regulatory element and a coding sequence is identified in a GWAS. (B) A CRE is removed with dual sgRNA/Cas. (C) The gene is mutated with the aid of NHEJ, resulting in a nonfunctional gene. (D) HDR is used to exchange one CRE variant to another, allowing for ecologically relevant insights via alternative allele swapping on different genomic backgrounds. (E) Similarly, HDR is used to alter a G to a C, potentially recapitulating a naturally occurring nonsynonymous amino acid change. Abbreviations: CDS, coding sequence; CRE, *cis*-regulatory element; CRISPR/Cas9, clustered regularly interspaced palindromic repeats-CRISPR-associated protein 9; DSB, double-strand break; GWAS, genome-wide association studies; HDR, homology-directed repair; NHEJ, non-homologous end-joining; sgRNA, single guide RNA.

From association to causality – choosing the right tools for the job

Different approaches for identifying candidate genes and/or mechanisms causing phenotypic variation necessitate different methods of functional validation. For example, RNAi is restricted to manipulation of whole transcript levels, and thus cannot be directly utilized to validate phenotypic differences caused by alternative alleles. By contrast, RNAi is a suitable tool to validate results from differential expression analysis because partial knockdown from RNAi can recapitulate differential expression patterns. In this sense, if an allele is associated with differential expression of a locus such as an expression (e)QTL or a candidate locus harboring variation only outside the coding region, RNAi can be used to investigate whether locus-specific expression levels produce a relevant phenotypic effect.

Box 3. Optimizing delivery methods for emerging model organisms

To address the issue of low throughput and/or difficulty of embryo injections, sgRNA/Cas riboprotein complexes has been tagged with a yolk protein precursor (YPP) peptide sequence and injected into the abdomen of female mosquitoes [60]. This method, termed receptor-mediated ovary transduction of cargo (ReMOT), circumvents the need for time-consuming embryo injections while still being sufficiently effective to produce mutant offspring in a reasonable timeframe. However, some major challenges are associated with the implementation of this method in other species. First, it relies upon a fusion between Cas protein and an ovary-targeting peptide that works efficiently in the focal species. Such a peptide (P2C) has been constructed for mosquitoes [60,90], but this is unlikely to work efficiently, or at all, outside higher dipterans [91]. Thus, testing whether the P2C peptide is transported into ovaries in the target species is a crucial step. Unfortunately, this requires custom synthesis of a fusion protein that brings together the ovary-targeting peptide and a fluorescent protein to assess transport [60,91,92]. If successful, a fusion of this ligand with Cas9 would need to be generated. The ReMOT approach also only induces mutations in the maternal lineage, such that heterozygous mutants in the progeny must have dominant phenotypes to enable screening of mutant progeny [60,91,92]. Finally, because novel genes are likely the focus of assessment, if knockout (KO) of such a gene is lethal, then neither high-penetrance mutations nor germline mutation is desirable, making this approach a problematic starting point for CRISPR/Cas9 gene manipulation in new species. Work to extend the use of ReMOT to enable transgenic replacement through homology-directed repair (HDR) is ongoing [60]. Excitingly, some recent work has been able to increase delivery efficiency using standard Cas9 proteins, thereby bypassing the need to create a species-specific peptide tag as in ReMOT. This new method, called direct parental (DIPA)-CRISPR [93], was successfully applied to cockroaches and red flour beetles, and can greatly facilitate the application of CRISPR in insects. Limitations include the need for staging vitellogenic females to maximize delivery success and the need to increase dramatically the amount of Cas9 and gRNA injected. Nonetheless, DIPA-CRISPR holds great promise, and researchers are encouraged to follow this line of research for future advances. Once working, the real challenge for the ecology and evolutionary biology community is to make gene editing using CRISPR-Cas relevant for addressing important questions in ecology and evolution (main text).

Depending upon the system, RNAi knockdown can be highly variable among treated individuals, resulting in quantitative trait variation that requires large sample sizes to achieve a significantly different phenotype compared to controls. In such a situation, the large effect of entire gene KO via CRISPR-Cas may be desirable.

The versatile toolbox of CRISPR-Cas enables validation at the level of genes, alleles, *cis*-regulatory elements, chromatin state, and methylation (discussed earlier). However, at present, the vast majority of studies on emerging model organisms have generated whole-gene KO mutations, and knock-in mutations have only been achieved in a few species to date to our knowledge [60–62]. Among the KO studies, these are dominated by somatic mutations in the coding region of genes affecting morphological phenotypes, which display a mosaic phenotype of wild-type and KO cell lineages, exemplified by the study of butterfly wing-color patterns where aberrant coloration phenotypes are readily visible within the same individual [18]. In addition, there are also some nice recent examples where the regulatory region has been successfully targeted to produce clear phenotypes [63–65].

For non-discrete phenotypes, germline transmission of the induced mutation enables the propagation of mutant lines, thus providing a route to large sample sizes of mutants for detailed comparison with controls. In addition, mutant lineages can be crossed with wild-type lineages to reduce the risk of off-target effects or to assess different genetic backgrounds, or can be crossed with other mutant lineages to create double mutants. However, such practices are to date uncommon in emerging model organisms. Instead, NHEJ-mediated disruption of the coding sequence is the most common way to use CRISPR-Cas, most likely because germline mutants are difficult to generate, especially because many candidate loci are core developmental genes whose germline KO is lethal (hence low-frequency somatic mosaic phenotypes are common). However, despite these CRISPR-Cas advances in emerging model species, it is important to keep in mind that such mosaic mutants rarely reflect ecologically relevant genetic variation (discussed in more detail below), and it is often difficult to know which genetic variants cause

the mosaicism because many alleles are generated, including mixes of homozygous and heterozygous clones, small and large deletions, and even null mutations.

Currently, the role of structural variation (SV; insertions, deletions, and inversions) in generating phenotypic variation remains an under-represented area of evolutionary genomics compared to SNP-focused studies [66], likely due to the fact that SVs can be much more challenging to identify than SNPs [67]. However, we note that, for deletions and inversions, ecologically relevant gene manipulation is readily available in the form of dual sgRNA-induced deletions. As an example, chromosomal inversions and deletions of large sequences have been achieved in zebrafish [68]. Similarly, a deletion of ~800 bp was induced with two separate DSB cuts in *Bombyx mori* [69]. It is likely that precise reconstitution of a naturally occurring deletion would require a large amount of screening to find mutants with exactly the same positions in the deletion, something which would be very important if the deletion occurs in a coding sequence. Nevertheless, ecologically relevant validation of SVs is now a possibility.

Interest in DNA methylation as a mechanism to generate phenotypic variation has dramatically increased over the past few years [70]. Although divergent DNA methylation patterns exists across populations [71] and species [72], evidence that variation in DNA methylation is the cause of adaptive phenotypic variation is scarce [73–75]. CRISPR-Cas-based technologies to sequence-specifically manipulate DNA methylation patterns could help to provide unequivocal tests for the role of methylation by ruling out potential confounding genetic effects (Box 4). Although such methods are still on the horizon for emerging model organisms, their implementation would contribute to a more thorough understanding of the increasingly studied, but still controversial, role of epigenetics in adaptation.

The challenges of functional validation in ecology and evolution

Despite the recent developments in CRISPR-Cas and RNAi for gene manipulation in emerging model organisms, challenges with these methods remain. These include low knock-in efficiencies, low spatial control of knockdown levels, and large amounts of time required per validated

Box 4. Manipulation of DNA methylation patterns associated with trait divergence

Interest in the role of epigenetic mechanisms in generating phenotypic variation has increased significantly in recent years [70–72,75,94,95]. This has been facilitated by technological advances in sequencing methods, for example chromatin immunoprecipitation and deep sequencing (ChIP-seq) for DNA-binding proteins [96] and bisulfite sequencing for 5-methylation of cytosine [97]. Of particular interest among evolutionary biologists has been DNA methylation, and this has been studied in several emerging model organisms [98].

Three different approaches can be used to validate the functionality of differentially methylated sites (DMS): inhibition of DNA methyltransferases with the use of drugs [75], targeted removal of the methylated sites (e.g., through knock-ins with CRISPR-Cas that remove CpG residues [99], or targeted methylation or demethylation [100] by using deactivated Cas variants fused to different epigenetic modifying enzymes. Although methyltransferase-inhibiting drugs are perhaps the easiest to apply, this strategy is also the most difficult to interpret because drugs alter DNA methylation genome-wide, thereby making it nearly impossible to dissect the direct causal effects, even if some level of tissue targeting can be done by tissue-specific injections [75].

Utilizing CRISPR-Cas to manipulate DNA methylation can be achieved in a blunt way by editing CpG residues through the HDR pathway [99]. This approach will also alter the genetic background, making interpretations potentially problematic. A preferred approach would be to use Cas variants in which the endonuclease functions have been abolished (catalytically dead Cas, dCas) that are fused to either DNA methyltransferases or DNA demethylases. These approaches rely on the sgRNA to guide the dCas to the region of interest where it invokes changes in methylation pattern depending on whether it carries a demethylase or a methyltransferase [101]. Such methods have been demonstrated to be applicable in several model systems (mainly in biomedical research [101]) but have to our knowledge not been performed in an emerging model organism. However, because CRISPR-Cas has a relatively low level of taxonomic restriction, emerging model organisms may play a significant role in deciphering the role of DNA methylation if CRISPR-Cas-mediated targeted methylation manipulation can be implemented. A major challenge concerns how to introduce the modified dCas and sgRNA into the cells or tissues of interest; this would presumably occur through integration of constructs into the genome through HDR.

locus. When weighing the investment of resources into functional validation against other venues of research (e.g., deeper sequencing, more significant *P* values for candidate loci), validation unfortunately often loses because it is perceived to be time-consuming and risky. Second, in situations when time and resources have been spent on functional validation of putative candidate genes with negative results, such results may be difficult to publish [76]. In this perspective, the ability of functional genetic methods to produce binary outcomes (locus is causative vs. not causative) rather than probabilistic evidence may instead be perceived as a curse rather than a cure. Regardless of whether negative results are (or are only perceived to be) difficult to publish, we argue that functional validation of loci is always of higher value than no validation at all, even when the results are negative. Rather than chastising our community for shying away from pursuing or requiring such validation, we argue that a well-documented negative RNAi/CRISPR-Cas result for robustly identified candidate locus would be fascinating, and such negative results are of importance to the community and the advancement of science [9].

Given that it is easier to detect and validate candidate loci underlying adaptive traits with a simple genetic basis (i.e., Mendelian or oligogenic traits), most advances in functional validation will occur for such phenotypes. This is especially likely when traits are morphological, discrete, and can be assessed using somatic mosaics.

A polygenic architecture poses several challenges when aiming to functionally verify genetic variants contributing to phenotypic variation. Because the effect size of individual loci is often very small for polygenic trait architectures, detecting the replacement of one allele with another via knock-in will require very large sample sizes to detect phenotypic effects. Polygenic traits might also be enriched for loci where gene-editing effects might be phenotypically masked by **genetic compensation** [77]. Unfortunately, these issues are likely to increase the already recognized bias in studies investigating the loci of adaptation towards traits with a simple genetic basis [8].

A significant challenge for the ecology and evolutionary biology community is to make functional validation using CRISPR-Cas ecologically relevant, here defined as gene editing that informs upon the phenotype in a way that reflects the natural variation in the candidate locus being investigated. At present the majority of published CRISPR-Cas studies in emerging model organisms have used a NHEJ-based approach for whole-gene KO (*cf* [63–65]). Although being potentially informative in mosaic phenotypes about the necessity of the gene for the formation of the investigated phenotype, these types of KO studies rarely, if ever, reconstitute the natural range or type of phenotypic variation being studied. There is a profound difference between removing the function of a locus (KO approach) compared to altering it from one functional naturally occurring variant to another (i.e., allelic variation).

In summary, there are two types of functional validation studies. In one type the researcher simply aims to validate the involvement of a gene in a trait, whereas in the other the researcher seeks to investigate whether a particular genetic variant has a different impact on a phenotype as compared to another variant, both of which are functional. We predict that increasing the efficiency of allelic replacements with CRISPR-Cas through the HDR pathway (Figure 1) will be an important factor in making functional validation with CRISPR-Cas ecologically relevant and therefore a truly powerful tool for the ecology and evolutionary biology community.

Concluding remarks

Functional validation studies are rapidly expanding from a few laboratory model species to a much greater diversity of organisms and traits in natural populations. This expansion, which

Outstanding questions

How can we increase the use of functional validation tools in the field of ecology and evolution to draw more robust inferences?

How can we increase the delivery efficiency and throughput of dsRNA and sgRNA/Cas for non-model species?

How can we move beyond the focus on simple Mendelian traits to also deploy functional genetic tools to investigate traits with poly- or omnigenic architectures?

Will relevant allele replacement (through the HDR pathway), as opposed to gene KO, impact on our understanding of the genetic basis of ecologically relevant traits?

Can we use CRISPR to experimentally induce SVs in the genome to examine their phenotypic consequences, and thus patterns of selection, to better understand the population genetic dynamics of SVs?

What is the relative role of epigenetic variation in phenotypic trait variation, and how can we facilitate the use of CRISPR to examine epigenetic variation?

provides a broader picture of how phenotypes develop and evolve, is driven by methodical developments such as more efficient and targeted delivery methods for RNAi and CRISPR-Cas (Box 3). Importantly, these developments are eroding the traditional divide between model and emerging model organisms, thereby facilitating a deeper understanding of the molecular underpinnings of phenotypic variation. We look forward to the development of more efficient HDR protocols for CRISPR-Cas gene editing to target both coding and noncoding regions of the genome because such developments have significant potential for making functional validation increasingly relevant to the study of ecology and evolutionary biology (see Outstanding questions).

Acknowledgments

We are grateful to Göran Arnqvist, David Berger, and Pascal Milesi, as well as to four anonymous reviewers and the editor, for comments that improved the manuscript. Figure 1 was adapted from 'CRISPR/Cas9 Gene Editing' by BioRender.com (2020), retrieved from https://app.biorender.com/biorender-templates. This study was supported by funding from the Swedish Research Council to A.H. (2020-03349), A.K. (2020-04298), and C.W.W. (2017-04386), and from the Agence National de la Recherche to A.K.

Declaration of interests

No interests are declared.

References

1. Vitti, J.J. *et al.* (2013) Detecting natural selection in genomic data. *Annu. Rev. Genet.* 47, 97–120
2. Jamann, T.M. *et al.* (2015) QTL mapping using high-throughput sequencing. In *Plant Functional Genomics: Methods and Protocols* (Alonso, J.M. and Stepanova, A.N., eds), pp. 257–285, Springer
3. van der Burg, K.R.L. *et al.* (2020) Genomic architecture of a genetically assimilated seasonal color pattern. *Science* 370, 721–725
4. Pardo-Diaz, C. *et al.* (2015) Towards the identification of the loci of adaptive evolution. *Methods Ecol. Evol.* 6, 445–464
5. Schielzeth, H. *et al.* (2018) Success and failure in replication of genotype–phenotype associations: How does replication help in understanding the genetic basis of phenotypic variation in outbred populations? *Mol. Ecol. Resour.* 18, 739–754
6. Liu, Y. *et al.* (2016) On the dependency of cellular protein levels on mRNA abundance. *Cell* 165, 535–550
7. Schielzeth, H. and Husby, A. (2014) Challenges and prospects in genome-wide quantitative trait loci mapping of standing genetic variation in natural populations. *Ann. N. Y. Acad. Sci.* 1320, 35–57
8. Wellenreuther, M. and Hansson, B. (2016) Detecting polygenic evolution: problems, pitfalls, and promises. *Trends Genet. TIG* 32, 155–164
9. Siddiq, M.A. *et al.* (2017) Experimental test and refutation of a classic case of molecular adaptation in *Drosophila* melanogaster. *Nat. Ecol. Evol.* 1, 25
10. Wu, M. *et al.* (2020) CRISPR/Cas9 mediated genetic resource for unknown kinase and phosphatase genes in *Drosophila*. *Sci. Rep.* 10, 7383
11. Verster, A.J. *et al.* (2014) Comparative RNAi screens in *C. elegans* and *C. briggsae* reveal the impact of developmental system drift on gene function. *PLoS Genet.* 10, e1004077
12. Tomczak, A. *et al.* (2018) Interpretation of biological experiments changes with evolution of the Gene Ontology and its annotations. *Sci. Rep.* 8, 5115
13. Stoeger, T. *et al.* (2018) Large-scale investigation of the reasons why potentially important genes are ignored. *PLoS Biol.* 16, e2006643
14. Klaassen, H. *et al.* (2018) CRISPR mutagenesis confirms the role of oca2 in melanin pigmentation in *Astyanax mexicanus*. *Dev. Biol.* 441, 313–318
15. Li, X. *et al.* (2015) Outbred genome sequencing and CRISPR/Cas9 gene editing in butterflies. *Nat. Commun.* 6, 8212
16. Mazo-Vargas, A. *et al.* (2017) Macroevolutionary shifts of WntA function potentiate butterfly wing-pattern diversity. *Proc. Natl. Acad. Sci. U. S. A.* 114, 10701–10706
17. Santos, M.E. *et al.* (2017) Taxon-restricted genes at the origin of a novel trait allowing access to a new environment. *Science* 358, 386–390
18. Woronik, A. *et al.* (2019) A transposable element insertion is associated with an alternative life history strategy. *Nat. Commun.* 10, 5757
19. Wucherpfennig, J.I. *et al.* (2019) Efficient CRISPR-Cas9 editing of major evolutionary loci in sticklebacks. *Evol. Ecol. Res.* 20, 107–132
20. Zhang, L. *et al.* (2017) Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence. *Proc. Natl. Acad. Sci. U. S. A.* 114, 10707–10712
21. Zhan, S. *et al.* (2014) The genetics of monarch butterfly migration and warning colouration. *Nature* 514, 317–321
22. Toubiana, W. and Khila, A. (2019) Fluctuating selection strength and intense male competition underlie variation and exaggeration of a water strider's male weapon. *Proc. Biol. Sci.* 286, 20182400
23. Wagner, G.P. and Lynch, V.J. (2010) Evolutionary novelties. *Curr. Biol.* 20, R48–R52
24. Casasa, S. *et al.* (2020) Nutrition-responsive gene expression and the developmental evolution of insect polyphenism. *Nat. Ecol. Evol.* 4, 970–978
25. Shabalina, S.A. and Koonin, E.V. (2008) Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.* 23, 578–587
26. Chapman, E.J. and Carrington, J.C. (2007) Specialization and evolution of endogenous small RNA pathways. *Nat. Rev. Genet.* 8, 884–896
27. Terenius, O. *et al.* (2012) RNA interference in Lepidoptera: an overview of successful and unsuccessful studies and implications for experimental design. *J. Insect Physiol.* 57, 231–245
28. Killiny, N. *et al.* (2014) Double-stranded RNA uptake through topical application, mediates silencing of five CYP4 genes and suppresses insecticide resistance in *Diaphorina citri*. *PLoS One* 9, e110536
29. Refki, P.N. *et al.* (2014) Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait. *Dev. Biol.* 392, 441–453
30. Xu, H.J. *et al.* (2015) Two insulin receptors determine alternative wing morphs in planthoppers. *Nature* 519, 464–467
31. Zhao, Y. *et al.* (2019) Mutations in NIInR1 affect normal growth and lifespan in the brown planthopper *Nilaparvata lugens*. *Insect Biochem. Mol. Biol.* 115, 103246
32. Zhang, C.-X. *et al.* (2019) Molecular mechanisms of wing polymorphism in insects. *Annu. Rev. Entomol.* 64, 297–314

33. Liu, Y. *et al.* (2016) Whole-animal genome-wide RNAi screen identifies networks regulating male germline stem cells in *Drosophila*. *Nat. Commun.* 7, 12149
34. Wang, G. *et al.* (2019) RNAi expression tuning, microfluidic screening, and genome recombineering for improved protein production in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 116, 9324–9332
35. Toubiana, W. *et al.* (2021) The growth factor BMP11 is required for the development and evolution of a male exaggerated weapon and its associated fighting behavior in a water strider. *PLoS Biol.* 19, e3001157
36. Feng, D. *et al.* (2019) RNA Interference by ingested dsRNA-expressing bacteria to study shell biosynthesis and pigmentation in *Crassostrea gigas*. *Mar. Biotechnol.* 21, 526–536
37. Whyard, S. *et al.* (2009) Ingested double-stranded RNAs can act as species-specific insecticides. *Insect Biochem. Mol. Biol.* 39, 824–832
38. Forsthoefel, D.J. *et al.* (2012) An RNAi screen reveals intestinal regulators of branching morphogenesis, differentiation, and stem cell proliferation in planarians. *Dev. Cell* 23, 691–704
39. Wang, Y. *et al.* (2011) Second-generation sequencing supply an effective way to screen RNAi targets in large scale for potential application in pest insect control. *PLoS One* 6, e18644
40. Vogel, E. *et al.* (2019) RNA interference in insects: protecting beneficials and controlling pests. *Front. Physiol.* 9, 1912
41. Schepers, U. (2006) *RNA Interference in Practice: Principles, Basics, and Methods for Gene Silencing in C. elegans, Drosophila, and Mammals*, John Wiley & Sons
42. Beach, L.Q. and Wade, J. (2015) Masculinisation of the zebra finch song system: roles of oestradiol and the z-chromosome gene tubulin-specific chaperone protein a. *J. Neuroendocrinol.* 27, 324–334
43. Beach, L.Q. *et al.* (2016) Inhibition of TrkB limits development of the zebra finch song system. *Brain Res.* 1642, 467–477
44. Haesler, S. *et al.* (2007) Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus area X. *PLoS Biol.* 5, e321
45. Podolska, K. and Svoboda, P. (2011) Targeting genes in living mammals by RNA interference. *Brief. Funct. Genomics* 10, 238–247
46. Raemdonck, K. *et al.* (2008) Maintaining the silence: reflections on long-term RNAi. *Drug Discov. Today* 13, 917–931
47. Smýkal, V. *et al.* (2020) Complex evolution of insect insulin receptors and homologous decoy receptors, and functional significance of their multiplicity. *Mol. Biol. Evol.* 37, 1775–1789
48. Gilles, A.F. and Averof, M. (2014) Functional genetics for all: engineered nucleases, CRISPR and the gene editing revolution. *EvoDevo* 5, 43
49. Jinek, M. *et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821
50. Perez-Pinera, P. *et al.* (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* 10, 973–976
51. Qi, L.S. *et al.* (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183
52. Vojta, A. *et al.* (2016) Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res.* 44, 5615–5628
53. Braun, S.M.G. *et al.* (2017) Rapid and reversible epigenome editing by endogenous chromatin regulators. *Nat. Commun.* 8, 560
54. Chen, B. *et al.* (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155, 1479–1491
55. Banta, J.A. and Richards, C.L. (2018) Quantitative epigenetics and evolution. *Heredity* 121, 210–224
56. Heryanto, C. *et al.* (2022) Mapping and CRISPR homology-directed repair of a recessive white eye mutation in *Plodia* moths. *iScience* 25, 103885
57. Bosch, J.A. *et al.* (2020) Gene knock-ins in *Drosophila* using homology-independent insertion of universal donor plasmids. *Genetics* 214, 75–89
58. Auer, T.O. *et al.* (2014) Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res.* 24, 142–153
59. Cui, Y. *et al.* (2018) Review of CRISPR/Cas9 sgRNA design tools. *Interdiscip. Sci. Comput. Life Sci.* 10, 455–465
60. Chaverra-Rodriguez, D. *et al.* (2018) Targeted delivery of CRISPR-Cas9 ribonucleoprotein into arthropod ovaries for heritable germline gene editing. *Nat. Commun.* 9, 3008
61. Wang, X. *et al.* (2020) CRISPR/Cas9 mediated ryanodine receptor I4790M knockin confers unequal resistance to diamides in *Plutella xylostella*. *Insect Biochem. Mol. Biol.* 125, 103453
62. Kistler, K.E. *et al.* (2015) Genome engineering with CRISPR-Cas9 in the mosquito *Aedes aegypti*. *Cell Rep.* 11, 51–60
63. Murugesan, S.N. *et al.* (2022) Butterfly eyespots evolved via cooption of an ancestral gene-regulatory network that also patterns antennae, legs, and wings. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2108661119
64. Lewis, J.J. *et al.* (2019) Parallel evolution of ancient, pleiotropic enhancers underlies butterfly wing pattern mimicry. *Proc. Natl. Acad. Sci. U. S. A.* 116, 24174–24183
65. Livraghi, L. *et al.* (2021) Cortex cis-regulatory switches establish scale colour identity and pattern diversity in *Heliconius*. *eLife* 10, e68549
66. Mérot, C. *et al.* (2020) A roadmap for understanding the evolutionary significance of structural genomic variation. *Trends Ecol. Evol.* 35, 561–572
67. Ho, S.S. *et al.* (2020) Structural variation in the sequencing era. *Nat. Rev. Genet.* 21, 171–189
68. Xiao, A. *et al.* (2013) Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* 41, e141
69. Ma, S. *et al.* (2012) Highly efficient and specific genome editing in silkworm using custom TALENs. *PLoS One* 7, e45035
70. Husby, A. (2022) Wild epigenetics – insights from epigenetic studies on natural populations. *Proc. Biol. Sci.* 298, 20211633
71. Heckwolf, M.J. *et al.* (2020) Two different epigenetic information channels in wild three-spined sticklebacks are involved in salinity adaptation. *Sci. Adv.* 6, eaaz1138
72. Bewick, A.J. *et al.* (2017) Evolution of DNA methylation across insects. *Mol. Biol. Evol.* 34, 654–665
73. Charlesworth, D. *et al.* (2017) The sources of adaptive variation. *Proc. Biol. Sci.* 284, 20162864
74. Cubas, P. *et al.* (1999) An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401, 157–161
75. Gore, A.V. *et al.* (2018) An epigenetic mechanism for cavefish eye degeneration. *Nat. Ecol. Evol.* 2, 1155–1160
76. Fanelli, D. (2012) Negative results are disappearing from most disciplines and countries. *Scientometrics* 90, 891–904
77. El-Brolosy, M. and Stainier, D. (2017) Genetic compensation: a phenomenon in search of mechanisms. *PLoS Genet.* 13, e1006780
78. Fire, A. *et al.* (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811
79. Hamilton, A.J. and Baulcombe, D.C. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286, 950–952
80. Wilson, R.C. and Doudna, J.A. (2013) Molecular mechanisms of RNA interference. *Annu. Rev. Biophys.* 42, 217–239
81. Bernstein, E. *et al.* (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363–366
82. Hammond, S.M. *et al.* (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293–296
83. Sun, D. *et al.* (2017) Progress and prospects of CRISPR/Cas systems in insects and other arthropods. *Front. Physiol.* 8, 608
84. Zhao, J. *et al.* (2019) Genome editing in large animals: current status and future prospects. *Natl. Sci. Rev.* 6, 402–420
85. Adli, M. (2018) The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* 9, 1911
86. Martin, A. *et al.* (2020) Bringing immersive science to undergraduate laboratory courses using CRISPR gene knockouts in frogs and butterflies. *J. Exp. Biol.* 223, jeb208793

87. Zhang, L. and Reed, R.D. (2017) A practical guide to CRISPR/Cas9 genome editing in Lepidoptera. In *Diversity and Evolution of Butterfly Wing Patterns: An Integrative Approach* (Sekimura, T. and Nijhout, H.F., eds), pp. 155–172, Springer
88. Li, M. *et al.* (2014) Efficient and heritable gene targeting in tilapia by CRISPR/Cas9. *Genetics* 197, 591–599
89. Kotwica-Rolinska, J. *et al.* (2019) CRISPR/Cas9 genome editing introduction and optimization in the non-model insect *Pyrrhocoris apterus*. *Front. Physiol.* 10, 891
90. Macias, V.M. *et al.* (2020) Cas9-mediated gene-editing in the malaria mosquito *Anopheles stephensi* by ReMOT control. *G3 (Bethesda)* 10, 1353–1360
91. Shirai, Y. and Daimon, T. (2020) Mutations in cardinal are responsible for the red-1 and peach eye color mutants of the red flour beetle *Tribolium castaneum*. *Biochem. Biophys. Res. Commun.* 529, 372–378
92. Heu, C.C. *et al.* (2020) CRISPR-Cas9-based genome editing in the silverleaf whitefly (*Bemisia tabaci*). *CRISPR J.* 3, 89–96
93. Shirai, Y. *et al.* (2022) DIPA-CRISPR is a simple and accessible method for insect gene editing. *Cell Rep. Methods* 2, 100215
94. Skinner, M.K. *et al.* (2014) Epigenetics and the evolution of Darwin's finches. *Genome Biol. Evol.* 6, 1972–1989
95. Viitaniemi, H.M. *et al.* (2019) Seasonal variation in genome-wide DNA Methylation patterns and the onset of seasonal timing of reproduction in great tits. *Genome Biol. Evol.* 11, 970–983
96. Barski, A. *et al.* (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823–837
97. Frommer, M. *et al.* (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. U. S. A.* 89, 1827–1831
98. Bonasio, R. (2015) The expanding epigenetic landscape of non-model organisms. *J. Exp. Biol.* 218, 114–122
99. Kang, J.G. *et al.* (2019) Regulation of gene expression by altered promoter methylation using a CRISPR/Cas9-mediated epigenetic editing system. *Sci. Rep.* 9, 11960
100. Devesa-Guerra, I. *et al.* (2020) DNA methylation editing by CRISPR-guided excision of 5-methylcytosine. *J. Mol. Biol.* 432, 2204–2216
101. Lei, Y. *et al.* (2018) DNA methylation and de-methylation using hybrid site-targeting proteins. *Genome Biol.* 19, 187