MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 3810–3822 doi: 10.1111/mec.13252

INVITED REVIEWS AND SYNTHESES

Connecting genotypes, phenotypes and fitness: harnessing the power of CRISPR/Cas9 genome editing

JEREMY M. BONO,* EUGENIA C. OLESNICKY* and LUCIANO M. MATZKIN†‡

*Department of Biology, University of Colorado Colorado Springs, 1420 Austin Bluffs Parkway, Colorado Springs, CO 80918, USA, †Department of Biological Sciences, University of Alabama in Huntsville, 301 Sparkman Drive, Huntsville, AL 35899, USA, ‡HudsonAlpha Institute for Biotechnology, 601 Genome Way, Huntsville, AL 35806, USA

Abstract

One of the fundamental goals in evolution and ecology is to identify the genetic basis of adaptive phenotypes. Unfortunately, progress towards this goal has been hampered by a lack of genetic tools available for nonmodel organisms. The exciting new development of the CRISPR (clustered regularly interspaced short palindromic repeat)/Cas9 (CRISPR-associated nuclease 9) genome-editing system now promises to transform the field of molecular ecology by providing a versatile toolkit for manipulating the genome of a wide variety of organisms. Here, we review the numerous applications of this groundbreaking technology and provide a practical guide to the creation of genetic knockouts, transgenics and other related forms of gene manipulation in nonmodel organisms. We also specifically discuss the potential uses of the CRISPR/Cas9 system in ecological and evolutionary studies, which will further advance the field towards the long-standing goal of connecting genotypes, phenotypes and fitness.

Keywords: genetic manipulation, genetics of adaptation, reverse genetics, RNA interference

Received 8 April 2015; revision received 20 May 2015; accepted 26 May 2015

Introduction

Understanding the genetic basis of adaptive traits has been a fundamental goal in ecology and evolutionary biology since the modern synthesis. This pursuit has gained considerable momentum over the last decade, fuelled in part by the rapid development of cost-effective *'-omics'* technologies, which have allowed access to the genome of nearly any organism (Feder & Mitchell-Olds 2003; Stapley *et al.* 2010; Pavey *et al.* 2012). While these technological breakthroughs have alleviated some of the challenges associated with identifying candidate genes linked to adaptive traits, molecular ecologists are now encountering another set of obstacles en route to the goal of uncovering the genetic basis of adaptation (Pavey *et al.* 2012). Many researchers find themselves with a list of candidate genes eager to take the next steps to advance their research programme. In principle, the road forward is obvious. Candidate genes need to be tested for links to phenotype, and ultimately, allelic variants must be examined for their influence on fitness (Dean & Thornton 2007; Piertney & Webster 2010; Storz & Wheat 2010; Barrett & Hoekstra 2011). Unfortunately, functional testing of candidate genes is most convincingly accomplished using genetic manipulations that involve the creation of genetic knockouts and transgenics, but such tools are not well developed in nonmodel organisms. The lack of readily available genetic toolkit for nonmodel organisms has served as a major roadblock that continues to impede progress towards understanding some of the key questions in ecology and evolution. The exciting news is that the recent development and rapid proliferation of genome-editing technologies, most notably the CRISPR/Cas9 system, promises to provide these tools, allowing investigators to more easily link patterns of variation between genotype, phenotype and fitness.

Our purpose in writing this review is to highlight the transformative potential of the CRISPR/Cas9 genome-editing system for the field of molecular ecology, and to provide a practical guide for researchers looking to get started using this technology with nonmodel organisms. Although the majority of studies utilizing the CRISPR/

Correspondence: Jeremy M. Bono, Fax: +1 719-255-3047; E-mail: jbono@uccs.edu

Cas9 system to date have not been performed in an explicit ecological context, we will review case studies, often from traditional model organisms, that demonstrate different applications of the technology that could be used in ecology and evolution. Drawing from these studies and our own experiences, we will highlight some of the considerations and challenges associated with the application of CRISPR/Cas9 to nonmodel organisms, along with potential strategies to circumvent these issues.

Introduction to the CRISPR/Cas9 genome-editing system

The CRISPR/Cas9 genome-editing system is derived from components of the microbial adaptive immune system that interact to protect against invasion by foreign nucleic acids (Ishino *et al.* 1987; Barrangou *et al.* 2007; Horvath & Barrangou 2010; Wiedenheft *et al.* 2012; Hsu *et al.* 2014). The most widely used type II CRISPR system of *Streptococcus pyogenes* involves three main components: the Cas9 endonuclease, *trans*-activating CRISPR RNA (tracRNA) and CRISPR RNAs (crRNA) (Deltcheva *et al.* 2011; Jinek *et al.* 2012). The crRNA includes sequence that is complementary to a foreign nucleic acid target, and after recruitment into the Cas9 complex by tracRNA, this complementary sequence guides Cas9 to the target site. For Cas9 to successfully bind the target, a sequence termed a protospacer adjacent motif (PAM), which serves to differentiate foreign DNA from the host genome, must be present just downstream of the target sequence. If present, Cas9 cleaves the foreign element just upstream of the PAM. The PAM required for Cas9 recognition in *S. pyogenes* is 5′-NGG, although the sequence varies for other CRISPR systems (Jinek *et al.* 2012; Esvelt *et al.* 2013; Hou *et al.* 2013).

For the purposes of genome editing, this system has been simplified by combining the tracRNA and crRNA into a single-guide RNA (sgRNA) that includes a custom-designed 18- to 20-bp sequence complementary to the desired target (Jinek *et al.* 2012) (Fig. 1A). These

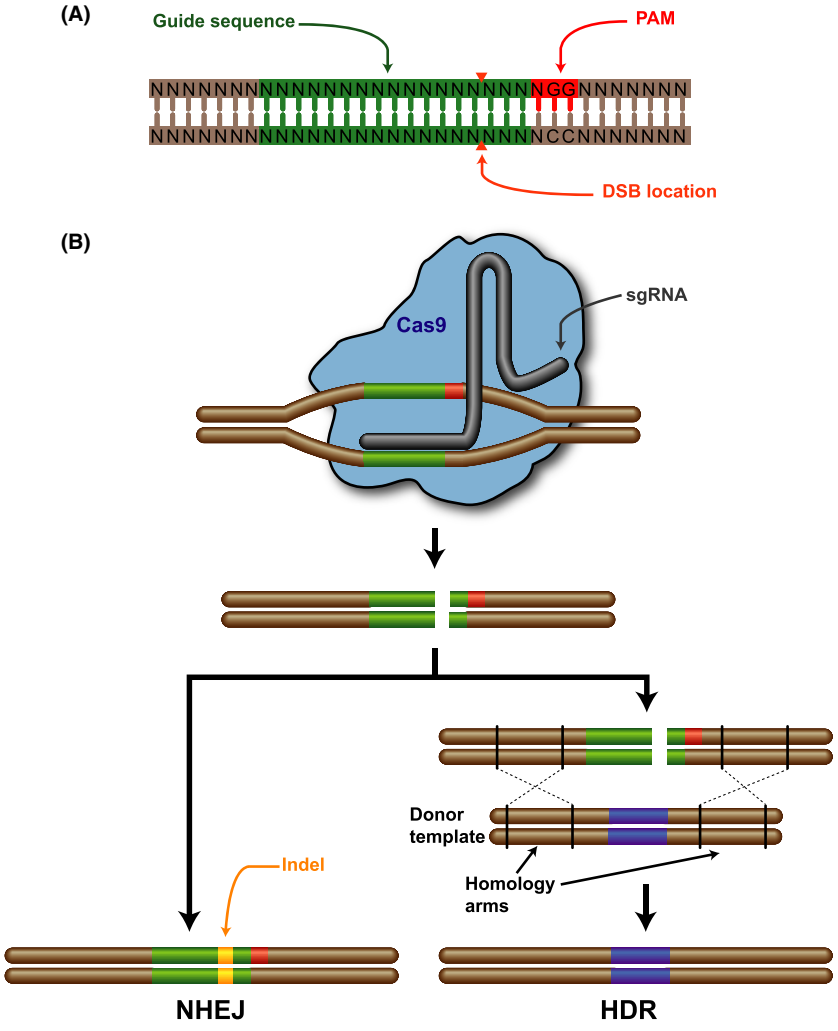

Fig. 1 (A) Target sequence including the complementary sequence of the crRNA or sgRNA (green), the PAM sequence (red) and the location of the double-stranded break (DSB). (B) Overview of the repair pathways that can be exploited for genome-editing purposes. The imprecise nonhomologous end joining (NHEJ) pathway often results in insertions and deletions, while homology-directed repair (HDR) pathway can be used to insert desired alterations into the genome. The green and red segments represent the sgRNA-binding and PAM sequences, respectively. The blue segment indicates the new recombined sequence.

components are typically delivered to early-stage embryos as *in vitro* transcribed RNA, or DNA carried within suitable expression vectors. When paired together with the sgRNA in the embryo, Cas9 will generate a double-strand break (DSB) at the targeted locus. Cells may then attempt to repair the DSB using one of two mechanisms. The most common is nonhomologous end joining (NHEJ), which is often imprecise, resulting in small insertions or deletions near the target site (Fig. 1B). Less frequently, cells may use homology-directed repair (HDR) mechanisms that utilize recombination to replace the affected region with a homologous sequence (Fig. 1B). The latter form of repair can be exploited to create transgenics with specific alterations to the targeted locus. Engineered homologous templates with desired sequence changes or sequence tags can easily be introduced along with the requisite CRISPR machinery to generate knock-ins, providing researchers with myriad tools for genetic manipulation and investigation of allelic variants. In either case, if the editing occurs in germ line cells, the mutation is potentially heritable and stable lines can thus be created.

Comparison with other reverse genetics approaches

In the past few decades, a variety of reverse genetic approaches have been developed to explore gene function and to link genetic variation to phenotypic differences within and between species. One fruitful approach involves the use of *in vitro* assays or a heterologous expression system to experimentally examine the functional consequences of variation in gene expression or protein structure. This strategy has been employed to identify causal mutations linked to a variety of phenotypes and has the potential to provide detailed resolution into the specific mechanisms of adaptation at the molecular level (Dean & Thornton 2007; Rebeiz & Williams 2011). A number of approaches have also been developed to assess gene function at the level of the whole organism, which allows for experiments that directly link genetic variation and fitness. This is likely where the CRISPR/Cas9 system will be most applicable and should prove superior to existing methods. In particular, antisense technologies such as RNA interference (RNAi) and morpholinos have been utilized to search for genes impacting specific phenotypes in many nonmodel organisms. While RNAi and morpholinos have revolutionized the fields of evolution and development and molecular ecology by allowing researchers to easily investigate gene function when it would otherwise not be possible, there are a number of caveats associated with these technologies. For example, due to the fact that both RNAi and morpholinos result in a reduction in targeted protein translation rather than a complete loss of function, in many instances a phenotype may not be detectable due to insufficient knockdown. Moreover, the effects are usually transient and highly variable among individuals. Consequently, these methods can produce a number of false-negative results. On the other hand, and more distressingly, many reports have uncovered a high incidence of off-target effects that lead to nonspecific phenotypes (Jackson *et al.* 2003, 2006; Lin *et al.* 2005; Birmingham *et al.* 2006; Kulkarni *et al.* 2006; Ma *et al.* 2006; Jackson & Linsley 2010; Kok *et al.* 2015). In fact, recent studies have shown that up to 80% of previously published phenotypes resulting from morpholino-mediated knockdown cannot be reproduced with genome-editing techniques such as CRISPR/Cas9, transcription activator-like effector nucleases (TALENs) or zinc-finger nucleases (ZFNs), suggesting that toxicity and off-target effects may produce the majority of these reported phenotypes (Kok *et al.* 2015). While the specificity of the CRISPR/Cas9 system is still under active investigation, a number of methods for reducing off-target effects have already been developed (Sander & Joung 2014) (see below for more complete discussion). The ease of use, promise of higher specificity, ability to permanently disable gene function, and the potential to create stable transgenics all combine to make the CRISPR/Cas9 a powerful alternative to antisense technologies.

The CRISPR/Cas9 system also compares favourably with other previously developed genome-editing systems such as ZFNs and TALENs. While these systems have been used in a number of model organisms to induce DSBs at specific targets, the CRISPR/Cas9 system is cheaper and much less laborious to implement, has higher efficiency and can be used to target multiple sites in the genome simultaneously simply by pairing Cas9 with more than one sgRNA (Harrison *et al.* 2014; Sander & Joung 2014). These practical considerations suggest that the CRISPR/Cas9 system is poised to supplant these technologies, as is already evidenced by the widespread adoption of CRISPR/Cas9 in numerous studies.

Applications of the CRISPR/Cas9 genome-editing system

The development of the CRISPR/Cas9 genome-editing system has already led to an explosion of studies in both model and nonmodel organisms since arriving on the scene only 2 years ago (Harrison *et al.* 2014; Sander & Joung 2014). Our goal for this section is not to provide an exhaustive review of these studies, but rather to use specific examples to highlight different applications of the technology that could be used to answer questions in ecology and evolution. We discuss the

examples along with recommendations for the types of experiments that could be conducted using the CRISPR/Cas9 system.

Exploiting the NHEJ pathway to create genetic knockouts and structural variants

Genetic knockouts are a classic tool for verifying functional links between genes and phenotypes. Attempts to repair DSBs induced by Cas9 using the NHEJ pathway often result in indels that can create diverse alleles, including loss-of-function alleles and hypomorphs. This method holds great promise for molecular ecologists working on nonmodel organisms, as simple experiments performed with genetic knockouts can verify links between a candidate gene and a specific phenotype associated with ecological adaptation (Fig. 2A). The creation of genetic knockouts has already been successful in a number of nonmodel species (Harrison *et al.* 2014). For example, Zhang *et al.* (2014) used CRISPR/Cas9 to study the function of *dmrt6* in Nile tilapia, *Oreochromis niloticus*. The *drmt* gene family includes a set of transcription factors involved in sex determination and differentiation in vertebrates. *Dmrt6* was previously thought to be present only in tetrapods, but a recent study reported an ortholog in coelacanth (Forconi *et al.* 2013), raising the possibility that orthologs may be found in other vertebrates as well. Zhang *et al.* (2014) found putative orthologs in Nile tilapia and several other fish species, suggesting a more ancient origin of *dmrt6* in bony fishes. Although high expression of *dmrt6* in the coelacanth testes suggested a role in reproduction, little was known about its specific function. Zhang *et al.* (2014) observed high expression in the testes of Nile tilapia, specifically after the initiation of spermatogenesis and appearance of spermatocytes. Given the spatial and temporal expression patterns, the researchers hypothesized that *dmrt6* may play a role in spermatogenesis. To examine this possibility, they used the CRISPR/Cas9 system to knock out *dmrt6* in developing embryos. Mutagenesis proved to be highly efficient, with 83% of fish injected with a combination of sgRNA and Cas9 mRNA incurring mutations. A subset of injected mutants was examined for several phenotypes associated with spermatogenesis. At 120 days post hatching, mutant fish had no spermatocytes, while controls had a normal number. In contrast, they found no difference in the amount of spermatogonia in mutant and control fish. Interestingly, mutants gradually recovered over time, reaching normal spermatocyte numbers and eventually producing viable sperm by 150 days post hatching. Thus, genetic knockout of *drmt6* appears to delay rather than permanently disable spermatogenesis. While Zhang *et al.* (2014) only systematically analysed the phenotype of fish that were injected as embryos, they also noted that mutations were detected in F_1 offspring, indicating that some of the induced mutations were heritable. Altogether, this study provided an early demonstration that the CRISPR/Cas9 system could be used to efficiently modify the genome of a nonmodel organism, thereby providing an important tool for linking genes and phenotypes.

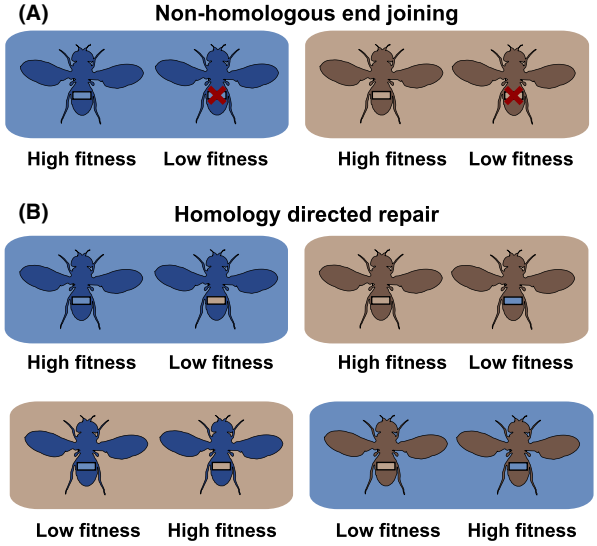

Fig. 2 Design and predictions for experiments using (A) genetic knockouts and (B) transgenics to assess the role of candidate loci in ecological adaptation. Each panel details an experiment using individuals from two populations adapted to divergent environments that are differentiated by colour (e.g. blue flies are adapted to the blue environment). The boxes inside the flies represent the status of the manipulated gene, with the colour reflecting the population of origin of each allele. The red X denotes a genetic knockout.

One of the advantages of the CRISPR/Cas9 system over other methods of genome modification is the potential for extensive multiplexing, which can be achieved by pairing multiple sgRNAs with Cas9 to simultaneously target different loci. Peng *et al.* (2015) provide a unique demonstration of this approach in the protozoan parasite *Trypanosoma cruzi*, the causal agent of Chagas' disease in humans. Despite the high impact this disease has on humans and other animals, research progress at the molecular level has been hampered by the complexity of the *T. cruzi* genome and the relative lack of tools for genetic manipulation. One of the major challenges stems from the unusually high number of large multigene families in *T. cruzi*, including several families whose members likely interact directly with immunological targets in hosts (Weston *et al.* 1999; El-Sayed *et al.* 2005; De Pablos & Osuna 2012). To assess the utility of the CRISPR/Cas9 system in this context,

the authors sought to simultaneously knockout multiple genes belonging to the 65 member β-galactofuronal glycosyltransferase (β-GalGT) gene family. High nucleotide sequence homology (93%) among members enabled Peng *et al.* (2015) to target the entire gene family with only three sgRNAs. Subsequent whole-genome sequencing revealed mutations in 63% of the β-GalGT genes, and failed to detect mutations at predicted off-target sites. The authors conclude that in the near future the CRISPR/Cas9 system will allow researchers to target many of the large multigene families that appear to interact directly with mammalian hosts. While this study used a limited number of sgRNAs to target many loci that share considerable homology, the system also allows simultaneous targeting of many unrelated loci with unique sgRNAs designed for each (e.g. Cong *et al.* 2013; Liu *et al.* 2014; Sakuma *et al.* 2014).

Although this study highlights the exciting multiplexing possibilities with the CRISPR/Cas9 system, it also illuminates some potential pitfalls that researchers may encounter when candidate genes belong to multigene families, or when the study organism is polyploid. The first complication arises due to the issue of functional redundancy. If different genes or homeologs (copies created through chromosomal duplication) have similar functions, then knocking out a single gene/copy may not cause a detectable phenotypic change. In such cases, the strategy employed by Peng *et al.* (2015) of knocking out multiple members of a family could be considered as a possible solution. In fact, a similar strategy was also recently adopted in rice (*Oryza sativa*) to knock out three members of the CDK gene family using one sgRNA (Endo *et al.* 2015). However, this also points to the other major complication that arises when working with gene families, paralogs or homeologs. Due to the possibility of unintended off-target effects, in situations where knocking out a single target is desired, extra precaution is warranted if the target belongs to a multigene family or the species is polyploid (see below for discussion of strategies to limit off-target effects).

In addition to exploiting the NHEJ pathway to create genetic knockouts, researchers have also used NHEJ to create precisely targeted structural modifications such as inversions and duplications (Blasco *et al.* 2014; Choi & Meyerson 2014; Maddalo *et al.* 2014; Torres *et al.* 2014). Kraft *et al.* (2015) performed a series of experiments in mice using two sgRNAs spaced varying distances apart on a chromosome to create structural modifications encompassing regions as large as 1.6 Mb. Specifically, imprecise repair triggered by the DSBs resulted in large deletions and inversions in 5% of cases, while duplications arose in 3% of cases. The authors note that the precision and efficiency of creating structural variations using the CRISPR/Cas9 system is much higher and also less laborious than previous methods (e.g. radiation).

The creation of precise structural variants may prove highly useful in evolution and ecology given the frequent involvement of chromosomal rearrangements and duplications in adaptation and speciation (Noor *et al.* 2001; Rieseberg 2001; Coghlan *et al.* 2005). For example, CRISPR/Cas9 may open the door to experiments involving direct manipulation of structural variation, which has not been a viable option in most systems. Furthermore, the creation of targeted structural variants may have other practical applications for molecular ecologists. As an example, fine mapping of quantitative trait loci (QTL) that lie within inversions is typically hindered by a lack of recombination, a problem that could potentially be overcome by targeting the genomic region to create colinear chromosomes. Targeted inversions might also be used to create balancer chromosomes, which have proven highly useful in some genetic model organisms as a tool for screening and maintaining mutant stocks that are sterile or lethal in the homozygous state (St Johnston 2002).

Exploiting the homology-directed repair mechanism to create transgenics

Another major application of the CRISPR/Cas9 system involves the creation of transgenics by introducing precise genetic modifications into the genome through co-option of the HDR pathway. When a donor template containing homology to the targeted region is delivered along with other CRISPR/Cas9 components, cells may use the HDR pathway to repair the DSB. Although this occurs much less frequently than repair with NHEJ, it is nonetheless becoming a well-established method for creating transgenics, with edits ranging in size from a single base pair to entire genes. Furthermore, some studies have shown that the efficiency of HDR can be improved by disrupting the competing NHEJ pathway (Beumer *et al.* 2013; Gratz *et al.* 2014; Chu *et al.* 2015; Maruyama *et al.* 2015).

Irion *et al.* (2014) provide a nice example of how the CRISPR/Cas9 system can be used in conjunction with repair templates to generate precise single-base modifications in zebrafish. Prior to this study, attempts at targeted knock-ins in zebrafish suffered from low efficiency, and the propagation of germ line mutations had not been documented. Irion *et al.* (2014) used the *albino (alb)* zebrafish mutant to assess the feasibility of generating knock-ins with the CRISPR/Cas9 system. Melanophores of *alb* mutant larvae are unable to produce melanin, and hence appear pale. This results from a single G->T mutation in exon 6 of *alb* that introduces a premature stop codon. Irion *et al.* (2014) used the

CRISPR/Cas9 system together with a donor template to repair the single-base mutation, which was easily assessed by screening larvae for pigmentation. After optimizing the design of donor templates, they were able to repair the mutation in nearly 50% of injected individuals. Moreover, a small percentage of these larvae transmitted the repaired allele to progeny, making this the first successful case of functional knock-in in the germ line of zebrafish. It is important to note that while the goal of this study was to engineer a small single-base modification, much larger changes are also possible.

Precise modification of target loci using the CRISPR/Cas9 system has several potential important applications in the field of molecular ecology. One of the major roadblocks in studying the genetic basis of adaptation is that often the fitness effects of alternative alleles need to be examined in different environments (Dean & Thornton 2007; Piertney & Webster 2010; Storz & Wheat 2010; Barrett & Hoekstra 2011). For some phenotypes, a possible solution is to examine functional differences between alleles using *in vitro* experiments (Dean & Thornton 2007). However, in cases where researchers wish to examine the phenotypic variation at the organismal level, often in conjunction with fitness experiments, the absence of genetic tools that allow the creation of transgenics poses a serious limitation. As an illustration of how the CRISPR/Cas9 system may help researchers overcome these challenges, imagine the common scenario where populations are locally adapted to divergent environments, and candidate genes for enhanced performance in each environment have been identified. With genetic manipulations generated using the CRISPR/Cas9 system, critical experiments could be conducted that were previously challenging or perhaps not even possible. For example, HDR could be utilized to replace an allele from a focal population with the corresponding allele from the other population. If the allele were involved in adaptation to the native environment of the focal population, then a reduction in the fitness of transgenics would be expected (Fig. 2B). Likewise, the putative adaptive allele from the focal population could be moved into individuals of the other population, in which case transgenics should exhibit higher fitness upon exposure to the environment of the focal population (Fig. 2B). If alleles differed at multiple nucleotide positions, a series of modifications could even target individual SNPs, providing a mechanism to uncover the specific substitutions responsible for adaptive phenotypes and potentially the order of mutational events. The power of this strategy has been demonstrated by previous studies using model organisms or *in vitro* assays to examine the phenotypic effects of individual mutations either alone or in combination with other mutations (e.g. Dean & Thornton 2007; Rebeiz *et al.* 2009; Tufts *et al.* 2015). The simplicity of the CRISPR/Cas9 system should help to further extend this approach to *in vivo* studies of nonmodel organisms.

In addition to editing nucleotide sequences, the HDR pathway can also be used to create other types of modifications that may prove useful to molecular ecologists. For example, reporters or tags can be inserted at precise locations to allow for visual tracking of transcripts or proteins *in vivo*. This tool could be especially useful when adaptive phenotypes involve changes in the pattern of gene expression. In such cases, visual reporters can be generated to determine the precise temporal and spatial differences among allelic variants or between orthologs of different species. While researchers have classically used antibody staining to analyse protein expression patterns *in situ*, antibodies for specific gene products are often unavailable for nonmodel organisms. Moreover, antibody generation and subsequent troubleshooting can be laborious, and commercial generation of synthetic or peptide antibodies is often unreliable and expensive. Thus, the ease of CRISPR/Cas9-mediated generation of knock-ins and tagged transgenics opens many new possibilities for precisely examining changes in gene expression. As proof-of-principle in the nonmodel mollusc, *Crepidula fornicata*, Perry & Henry (2015) created transgenic embryos expressing the mCherry reporter fused to the gene β-*catenin*. The reporter allowed them to observe the expression pattern of β-*catenin* during embryonic development, which suggested a role in cell adhesion and signal transduction.

Using the CRISPR/Cas9 system to regulate gene expression

Aside from generating knockouts and transgenics, there are also some other exciting applications of the CRISPR/Cas9 system in the development pipeline. Most notably, a catalytically inactive version of Cas9 ('dead' Cas9; dCas9) has been used to regulate endogenous gene expression by either repressing or activating genes of interest (Fig. 3) (Gilbert *et al.* 2013; Qi *et al.* 2013). dCas9 lacks endonuclease activity but is still able to bind sequences targeted by sgRNA. CRISPR interference (CRISPRi) involves recruitment of dCas9 to promoter regions or areas just downstream from transcription start sites, which effectively blocks transcription. This has been further optimized by fusing dCas9 to a repressor domain, which seems to result in more efficient knockdown (Gilbert *et al.* 2013, 2014). This system is similar in its effects to RNAi, but may offer advantages in terms of greater specificity and fewer off-target effects. In addition, not all species have the

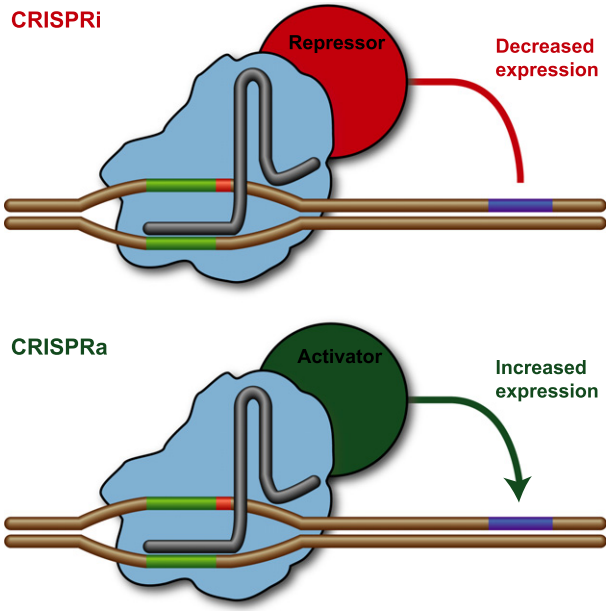

Fig. 3 A catalytically inactive version of Cas9 (dead or dCas9) can be used to activate (CRISPRa) or repress gene expression (CRISPRi).

required cellular machinery for RNAi to work, making CRISPRi an attractive alternative. dCas9 has also been used to activate gene transcription (CRISPRa), by recruiting multiple activating effector domains that can initiate transcription (Gilbert *et al.* 2014).

Manipulation of gene expression would be quite relevant when differences between divergently adapted populations are transcriptional in nature. Moreover, the ability to knock down or activate gene expression may be important when the generation of a full loss-of-function allele is sterile or lethal. In such cases, mutant alleles will not be useful for studying adaptive traits, or permanent lines may not be possible to maintain. However, gene knockdown or overexpression may circumvent these problems and concomitantly provide clues to gene function. While CRISPRi and CRISPRa have currently been implemented in only a few study systems, these strategies have already been optimized to modulate the expression of approximately 16 000 human genes, as well as multiple bacterial and plant genes without any detectable toxicity (Gilbert *et al.* 2013, 2014; Piatek *et al.* 2014). We expect that continued development of CRISPRa and CRISPRi will see applications in a variety of organisms in the near future.

A practical guide to using CRISPR/Cas9 in nonmodel organisms

The CRISPR/Cas9 system has the potential to revolutionize many research fields, including molecular ecology. As such, we expect molecular ecologists will be eager to determine whether the technology is transferable to their study system. In this section, we discuss some of the important considerations researchers will need to think about when adopting this technology, in particular highlighting some of the challenges and limitations that may arise when working with nonmodel organisms.

Perhaps the strongest limiting factor in adopting CRISPR/Cas9 technology in nonmodel organisms will be the feasibility of delivering the necessary components to early-stage embryos. Many studies in ecology and evolution require relatively large sample sizes, and given that efficiency of mutagenesis can be quite low, this means researchers will likely need easy access to a large number of developing embryos. Assuming this is possible, the CRISPR/Cas9 components then need to be delivered to embryos early in development, which is typically accomplished through microinjection, electroporation, or infection via virus or bacteria. While these delivery mechanisms are routinely used in laboratories working on model organisms, protocols can be technically challenging and may not be immediately transferable to nonmodel systems. This is one area where collaboration with laboratories working on model organisms will prove highly beneficial. In addition, several commercial vendors do offer microinjection services and may be willing to adapt protocols to other species as demand increases. While optimization of delivery methods will undoubtedly pose a challenge, the variety of organisms for which successful protocols have already been developed is encouraging (Harrison *et al.* 2014).

The other main limitation in adopting CRISPR technology concerns whether mutant lines can be maintained and propagated. In some cases, it may be possible to perform experiments on the F_0 generation, as, for example, in the study on Nile tilapia described above (Zhang *et al.* 2014). However, F_0 mutants will often be genetic mosaics, with different cells carrying alternative mutations (or no mutations), which could prove problematic depending on the goals of the study. In many cases, stable mutant lines will need to be established and maintained for subsequent experimentation, making species that are easily reared in the laboratory ideal candidates for CRISPR/Cas9 experiments.

Strategies for constructing sgRNAs

CRISPR/Cas9 components can be delivered to embryos as RNA or DNA carried on an expression vector (Box 1). At this point, there is no clear consensus on which method results in higher efficiency, and this may also vary for different study species. While a consensus

Box 1. Design and delivery of CRISPR/Cas9 components

CRISPR/Cas9 components can be easily prepared in laboratories using routine molecular protocols and equipment readily available to most molecular ecologists. Typically, Cas9 and sgRNA are delivered to embryos either as DNA carried on expression vectors or as RNA. Expression vectors can be purchased commercially, or from Addgene (www.addgene.org/CRISPR/), an online plasmid repository. Although some researchers prefer to work with DNA rather than RNA due to potential problems with RNA stability, it should also be noted that available expression vectors may not work efficiently across species due to divergence in promoter regions. Hence, new plasmids carrying species-specific promoters may need to be generated, which requires knowledge of the promoter sequence and additional labour. In such cases, delivery of mRNA may prove to be a simpler solution.

Cas9: Cas9 expression vectors can be directly delivered to embryos. Numerous vectors are already available, including some designed with promoters that allow for inducible or tissue-specific expression of Cas9. Tissue-specific or inducible expression of Cas9 may allow for the investigation of genes that would otherwise cause early embryonic lethality or other undesired pleiotropic effects. Cas9 expression vectors can also be used to generate mRNA through *in vitro* transcription. If using this method, a 5′ cap and polyadenylated tail must be added to the mRNA, which can be accomplished using commercially available kits. Ready-to-use Cas9 mRNA can also be purchased from commercial vendors.

Guide RNA: sgRNA is designed to be complementary to an 18- to 20-base target sequence on either DNA strand. The only requirement is that the sequence lies immediately 5′ of the PAM (NGG) motif, which is not included as part of the sgRNA. Online tools can be used to evaluate potential targets and optimize design of sgRNAs, taking into consideration sequence features that are predicted to improve efficiency (e.g. Doench *et al.* 2014). If the guide is to be delivered as DNA, plasmids can be purchased that allow for one or more sgRNAs to be cloned into the vector. Generally, this involves ligation of two annealed oligos containing the corresponding target sequence into the vector (protocols will vary depending on the plasmid), which can be delivered to embryos following verification and clean-up using commercially available kits.

Alternatively, the sgRNA can also be delivered as RNA. A simple protocol to make the guide involves a polymerase chain reaction (PCR) amplification step to generate a DNA template, which is then used for *in vitro* transcription (e.g. Bassett *et al.* 2013). PCR uses a common reverse primer containing sequence corresponding to essential portion of crRNA and tracRNA of the sgRNA (5′-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTT**GCTATTTCTAGCTCTAAAAC**-3′). The custom forward primer (5′-GAAATTAATACGACTCACTATA*GG*$N_{18\text{-}20}$**GTTTTAGAGCTAGAAATACG**-3′) contains a T7 promoter-binding sequence (underlined), and the 18- to 20-nt target sequence followed by an overlapping complement of the reverse primer (in bold). The *GG* is required for efficient transcription by T7 polymerase; it can be part of the target sequence or appended as extra nucleotides. These two primers are used in a no-template PCR to generate a double-stranded product that is then used for *in vitro* transcription. The RNA produced by the IVT reaction can then be delivered to embryos following clean-up using commercially available kits.

might eventually emerge as more systematic studies on multiple species are conducted, it is important to note that both methods have been used with reasonably high efficiency in a number of different organisms. Regardless of which method is chosen, the first step is to synthesize the sgRNA (Box 1). The sgRNA can be designed to target essentially any region, provided that it lies immediately proximal to a PAM sequence ('NGG'). Assuming only one locus is being targeted, the simplest scenario involves the use of one sgRNA in conjunction with Cas9. While the use of one sgRNA has proven to be successful in a number of studies, there are reasons to consider using multiple sgRNAs even when only targeting a single locus. First, studies have demonstrated that sgRNAs often vary considerably in efficiency, but, as of now, the reasons for this are not clear (Ran *et al.* 2013). As delivery of CRISPR/Cas9 components to embryos is much more laborious and challenging than making sgRNAs, it might make sense to use more than one sgRNA to increase the overall success rate. Another reason to consider using multiple sgRNAs is for screening purposes (Box 2). When two DSBs occur, repair by NHEJ often results in deletion of the entire region separating the breaks, meaning that the mutation is generally much larger than most induced by a single sgRNA (e.g. Paix *et al.* 2014). These larger mutations are more easily detected during the screening process, which can lead to considerable cost savings. One drawback of using multiple sgRNAs is that it might increase the probability of unintended off-target effects. However,

Box 2. Methods to screen for mutations

Although in some cases it may be possible to perform experiments with the injected generation (F_0), often researchers will want to create stable homozygous mutant lines to be used for later experiments. This requires a crossing scheme in conjunction with a mechanism to screen large numbers of individuals for mutations (assuming there is no visible marker). Although the injected generation can be screened for mutations, it is important to note that in most cases these individuals will be genetic mosaics, with potentially many different mutations per individual. Although screening is still possible, it could be messy given the possibility of a large number of alleles, and any mutant alleles that are detected may not be from the germ line. Regardless of whether the injected generation is screened for mutations, the first step will generally be to backcross these individuals to wild type (this can be performed for multiple generations to reduce off-target effects). The progeny of these backcrosses will need to be screened for mutations, which involves genotyping a large number of individuals over multiple generations. Below, we outline several different genotyping options, considering potential trade-offs between time, detection efficiency and cost.

Sequencing: Sequencing is a highly sensitive method for detecting mutations. Although a mutation will at some point need to be sequenced for verification, this is generally not very cost-effective for screening large numbers of individuals.

High-resolution melt analysis (HRMA): HRMA provides a rapid, highly efficient method for detecting mutations. It is powerful enough to detect even single-base-pair differences between samples, so it could be used to detect indels generated by imprecise NHEJ or SNPs resulting from HDR. While probably the fastest method for genotyping and also very reliable, HRMA can carry significant start-up costs depending on what equipment is available. The method requires a real-time qPCR instrument and may also require the purchase of additional software to perform HRMA.

Screening by PCR amplicon size: Most indels created using only one sgRNA are likely to be smaller than can be easily detected on an agarose gel. However, multiple sgRNAs simultaneously targeting different regions of the locus can be used to generate larger deletions, because repair often involves joining together the two cut sites. PCR amplicons with large indels could thus be easily detected on an agarose gel. This would provide a highly cost-effective, relatively rapid method of screening. However, individuals with mutations occurring at only one of the cut sites might be overlooked as indels are likely to be small. Smaller indels can also be detected using capillary electrophoresis usually in conjunction with fluorescence detection (e.g. as is used for microsatellites). Once fluorescent primers are purchased, fragment analysis can be very cost-effective because fluorescent labelling and the ability to design primers to amplify fragments of varying sizes permit a high degree of multiplexing.

Nuclease assays: Several commercially available nucleases (e.g. T7 endonuclease, CEL I, Surveyor nuclease, or mung bean nuclease) recognize and cleave mismatches between DNA strands due to single nucleotide polymorphisms or small indels (Oleykowski *et al.* 1998; Babon *et al.* 2003; Qiu *et al.* 2004; Till *et al.* 2004). Digestion products can then be separated using standard gel electrophoresis methods.

Restriction fragment length polymorphism: sgRNAs can be designed to include a restriction site, such that a deletion will likely result in a loss of the site. This can be cost-effective but may not detect all mutations depending on the exact location of the restriction site, and could severely limit the number of potential targets. This method would be particularly useful for experiments where alleles are exchanged between populations using HDR, assuming that there are restriction site differences between the alleles.

there are a number of strategies that can be easily implemented to reduce this possibility (see below for more discussion).

Design of donor repair templates

For experiments designed to generate transgenics via HDR, a donor repair template containing the desired edits must be provided along with other components to trigger homologous recombination. Repair templates can be double-stranded DNA cloned into plasmids or, in the case of smaller modifications (e.g. ≤200 bp), single-stranded oligonucleotides (ssODNs). To facilitate recombination, the repair template must include homology arms that flank the sequence to be inserted. At this point, the optimal length of flanking homology arms is not clear, and conflicting reports from different studies likely, at least in part, reflect variability across study systems. As a general guideline, for eukaryotes 200–400 bases of homology is often considered the minimum for efficient recombination using large repair templates, with optimal efficiency typically achieved at longer lengths (Hasty *et al.* 1991; Waldman 2008; Wu *et al.* 2008; Beumer *et al.* 2013; Shin *et al.* 2014). However, a

recent study in *C. elegans* demonstrated successful insertion of green fluorescent protein (864 bases) with considerably shorter homology arms (59/59-nt) (Paix *et al.* 2014). A much shorter region of homology has generally been used for ssODNs (Radecke *et al.* 2010; Bedell *et al.* 2012; Beumer *et al.* 2013; Chen *et al.* 2014; Paix *et al.* 2014). Paix *et al.* (2014) also demonstrated more efficient editing in *C. elegans* when homology arms immediately flanked the cut site rather than being placed even a relatively short distance away. For experiments described above involving the exchange of alleles between populations (Fig. 2B), the alleles themselves may provide enough homology to facilitate recombination without the need for additional homologous flanking sequence. However, it should be noted that previous research has demonstrated that recombination frequency declines with a reduction in sequence identity (Waldman 2008; Heyer *et al.* 2010), although this may also depend on the precise location of mismatches (Negritto *et al.* 1997). This could pose a problem in cases where alleles of different populations are considerably divergent. Despite some uncertainty concerning the optimal design of donor templates, gene targeting has been successful in many different study systems, suggesting that these challenges can be easily overcome.

Another consideration in the design of repair templates is that if the target sequence is present in the repair template, it may be advantageous to modify the sequence so that Cas9 does not also cut the template or the edited sequence following recombination. For example, the introduction of mutations to the PAM sequence, or other synonymous mutations in the target sequence can be used to disrupt recognition (Yang *et al.* 2014). This is straightforward if working with custom-synthesized ssODNs, but will require extra steps when using PCR-generated templates, as mutations will need to be introduced (e.g. through site-directed mutagenesis). In cases where exogenous sequence will be inserted (e.g. a fluorescent tag), the sgRNA can be designed so that the tag will disrupt the target sequence upon insertion.

Strategies for minimizing off-target effects

The targeting specificity of the CRISPR/Cas9 system is a subject of active debate and continued experimentation (Sander & Joung 2014). While the propensity to produce mutations at unintended target sites continues to be evaluated, it must be acknowledged that there will always be at least some risk of off-target mutagenesis and strategies to minimize this possibility should be considered. There are two basic ways to combat this issue. The first involves methods to enhance target specificity in order to prevent off-target effects in the first place, and the second involves experimental designs that minimize the confounding effects of any off-target mutagenesis that does happen to occur.

For organisms with available genome sequences, researchers can limit off-target effects by carefully designing the guide RNA and removing complementarity to potential off-target sequences within the genome. This can easily be accomplished using one of a growing number of publically available databases dedicated to generating unique sgRNAs. The 12-nt 'seed' region occurring proximal to the PAM site appears to be most important for specificity as mismatches in this region are less often tolerated, while mismatches in the rest of the target sequence are less problematic (Cong *et al.* 2013). Unfortunately, many molecular ecologists work on species that do not have sequenced genomes, which means that the uniqueness of target sites cannot be assessed. Nevertheless, several strategies have emerged that increase target specificity by altering either sgRNAs or Cas9. Fu *et al.* (2014) showed that truncated sgRNAs consisting of 17–18 nucleotides instead of 20 (tru-gRNAs) can dramatically reduce off-target cleavage without compromising efficiency. Likewise, adding two extra guanines to the 5′ end of the sgRNA (termed ggX_{20} sgRNA) also appears to increase specificity with no loss of efficiency (Cho *et al.* 2014; Kim *et al.* 2015). A 'dual nicking' strategy utilizing a mutant nickase version of Cas9 (Cas9n) that only produces a single-strand break (SSB) has also been shown to substantially reduce off-target effects (Ran *et al.* 2013). With this strategy, Cas9n is paired with two appropriately spaced sgRNAs targeting opposite strands of the genomic DNA. A DSB thus only occurs if both sgRNAs locate their targets, which effectively doubles the number of bases needed for recognition, thereby greatly improving specificity.

A number of experimental approaches can also be used to minimize the confounding effects of off-target mutations should they occur. One solution is to perform experiments with mutants created using more than one sgRNA, each specific to a different region of the target locus. Because the two sgRNAs would be unlikely to have the same off-target effects, concordance in phenotype in experimental assays on the different mutants would help to rule out other mutations as a potential cause. In established model organisms, the standard for 'cleaning up' mutagenized chromosomes is to backcross mutants to wild-type stocks multiple times before creating a homozygous mutant line, which is likely to break up an association between the desired mutation and any off-target mutations. If two lines carrying different mutations are created this way, they can also be crossed prior to the experiment to create trans-heterozygotes to

further reduce the likelihood of off-target effects. In fact, the use of trans-heterozygotes and confirmation of phenotypes among lines harbouring different mutant alleles, and the use of rescue experiments to confirm mutational specificity, remains the gold standard for genetic investigation in model organisms. Fortunately, the ease with which multiple mutations can be generated using the CRISPR/Cas9 system makes all of these alternatives highly feasible. Researchers may choose to use one or several of these strategies depending on particulars of the study system (e.g. generation time).

Conclusions

The development of the highly versatile CRISPR/Cas9 genome-editing system, in conjunction with continued advances in genomic technologies, promises to usher in an exciting new era of molecular ecology. Continued evaluation of the CRISPR/Cas9 system in different study organisms will undoubtedly lead to further improvements and additional applications in the near future. In particular, additional research on the potential for off-target effects is necessary, and this will likely motivate further enhancements to targeting specificity. With the necessary tools now available to connect genotypes, phenotypes and fitness in a wide variety of study systems, answers to many long-standing questions in ecology and evolution are within reach.

Acknowledgements

The authors would like to thank Darrell Killian and Meghan Lybecker for insightful discussions.

References

Babon JJ, McKenzie M, Cotton RGH (2003) The use of resolvases T4 endonuclease VII and T7 endonuclease I in mutation detection. *Molecular Biotechnology*, **23**, 73–81.

Barrangou R, Fremaux C, Deveau H *et al.* (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, **315**, 1709–1712.

Barrett RDH, Hoekstra HE (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews. Genetics*, **12**, 767–780.

Bassett AR, Tibbit C, Ponting CP, Liu J-L (2013) Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Reports*, **4**, 220–228.

Bedell VM, Wang Y, Campbell JM *et al.* (2012) In vivo genome editing using a high-efficiency TALEN system. *Nature*, **491**, 114–118.

Beumer KJ, Trautman JK, Mukherjee K, Carroll D (2013) Donor DNA utilization during gene targeting with zinc-finger nucleases. *G3 (Bethesda, MD)*, **3**, 657–664.

Birmingham A, Anderson EM, Reynolds A *et al.* (2006) 3′ UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nature Methods*, **3**, 199–204.

Blasco RB, Karaca E, Ambrogio C *et al.* (2014) Simple and rapid in vivo generation of chromosomal rearrangements using CRISPR/Cas9 technology. *Cell Reports*, **9**, 1219–1227.

Chen L, Tang L, Xiang H *et al.* (2014) Advances in genome editing technology and its promising application in evolutionary and ecological studies. *GigaScience*, **3**, 24.

Cho SW, Kim S, Kim Y *et al.* (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Research*, **24**, 132–141.

Choi PS, Meyerson M (2014) Targeted genomic rearrangements using CRISPR/Cas technology. *Nature Communications*, **5**, 3728.

Chu VT, Weber T, Wefers B *et al.* (2015) Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology*, **33**, 543–548.

Coghlan A, Eichler EE, Oliver SG, Paterson AH, Stein L (2005) Chromosome evolution in eukaryotes: a multi-kingdom perspective. *Trends in Genetics: TIG*, **21**, 673–682.

Cong L, Ran FA, Cox D *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, NY)*, **339**, 819–823.

De Pablos LM, Osuna A (2012) Multigene families in *Trypanosoma cruzi* and their role in infectivity. *Infection and Immunity*, **80**, 2258–2264.

Dean AM, Thornton JW (2007) Mechanistic approaches to the study of evolution: the functional synthesis. *Nature Reviews. Genetics*, **8**, 675–688.

Deltcheva E, Chylinski K, Sharma CM *et al.* (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, **471**, 602–607.

Doench JG, Hartenian E, Graham DB *et al.* (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nature Biotechnology*, **32**, 1262–1267.

El-Sayed NM, Myler PJ, Bartholomeu DC *et al.* (2005) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science (New York, NY)*, **309**, 409–415.

Endo M, Mikami M, Toki S (2015) Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant & Cell Physiology*, **56**, 41–47.

Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nature Methods*, **10**, 1116–1121.

Feder ME, Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Nature Reviews. Genetics*, **4**, 651–657.

Forconi M, Canapa A, Barucca M *et al.* (2013) Characterization of sex determination and sex differentiation genes in Latimeria. *PLoS ONE*, **8**, e56006.

Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology*, **32**, 279–284.

Gilbert LA, Larson MH, Morsut L *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, **154**, 442–451.

Gilbert LA, Horlbeck MA, Adamson B *et al.* (2014) Genome-scale CRISPR-mediated control of gene repression and activation. *Cell*, **159**, 647–661.

Gratz SJ, Ukken FP, Rubinstein CD *et al.* (2014) Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics*, **196**, 961–971.

Harrison MM, Jenkins BV, O'Connor-Giles KM, Wildonger J (2014) A CRISPR view of development. *Genes & Development*, **28**, 1859–1872.

Hasty P, Rivera-Pérez J, Bradley A (1991) The length of homology required for gene targeting in embryonic stem cells. *Molecular and Cellular Biology*, **11**, 5586–5591.

Heyer W-D, Ehmsen KT, Liu J (2010) Regulation of homologous recombination in eukaryotes. *Annual Review of Genetics*, **44**, 113–139.

Horvath P, Barrangou R (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science (New York, NY)*, **327**, 167–170.

Hou Z, Zhang Y, Propson NE *et al.* (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 15644–15649.

Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, **157**, 1262–1278.

Irion U, Krauss J, Nüsslein-Volhard C (2014) Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. *Development (Cambridge, United Kingdom)*, **141**, 4827–4830.

Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, **169**, 5429–5433.

Jackson AL, Linsley PS (2010) Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nature Reviews Drug Discovery*, **9**, 57–67.

Jackson AL, Bartz SR, Schelter J *et al.* (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnology*, **21**, 635–637.

Jackson AL, Burchard J, Schelter J *et al.* (2006) Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA (New York, NY)*, **12**, 1179–1187.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, NY)*, **337**, 816–821.

Kim D, Bae S, Park J *et al.* (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nature Methods*, **12**, 237–243.

Kok FO, Shin M, Ni C-W *et al.* (2015) Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Developmental Cell*, **32**, 97–108.

Kraft K, Geuer S, Will AJ *et al.* (2015) Deletions, inversions, duplications: engineering of structural variants using CRISPR/Cas in mice. *Cell Reports*, **10**, 833–839.

Kulkarni MM, Booker M, Silver SJ *et al.* (2006) Evidence of off-target effects associated with long dsRNAs in *Drosophila melanogaster* cell-based assays. *Nat Meth*, **33**, 833833–838838.

Lin X, Ruan X, Anderson MG *et al.* (2005) siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. *Nucleic Acids Research*, **33**, 4527–4535.

Liu Y, Ma S, Wang X *et al.* (2014) Highly efficient multiplex targeted mutagenesis and genomic structure variation in *Bombyx mori* cells using CRISPR/Cas9. *Insect Biochemistry and Molecular Biology*, **49**, 35–42.

Ma Y, Creanga A, Lum L, Beachy PA (2006) Prevalence of off-target effects in *Drosophila* RNA interference screens. *Nature*, **443**, 359–363.

Maddalo D, Manchado E, Concepcion CP *et al.* (2014) In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature*, **516**, 423–427.

Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL (2015) Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nature Biotechnology*, **33**, 538–542.

Negritto MT, Wu X, Kuo T, Chu S, Bailis AM (1997) Influence of DNA sequence identity on efficiency of targeted gene replacement. *Molecular and Cellular Biology*, **17**, 278–286.

Noor MA, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences*, **98**, 12084–12088.

Oleykowski CA, Bronson Mullins CR, Godwin AK, Yeung AT (1998) Mutation detection using a novel plant endonuclease. *Nucleic Acids Research*, **26**, 4597–4602.

Paix A, Wang Y, Smith HE *et al.* (2014) Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 Sites in *Caenorhabditis elegans*. *Genetics*, **198**, 1347–1356.

Pavey SA, Bernatchez L, Aubin-Horth N, Landry CR (2012) What is needed for next-generation ecological and evolutionary genomics? *Trends in Ecology & Evolution (Personal Edition)*, **27**, 673–678.

Peng D, Kurup SP, Yao PY, Minning TA, Tarleton RL (2015) CRISPR-Cas9-mediated single-gene and gene family disruption in *Trypanosoma cruzi*. *MBio*, **6**, e02097-2014.

Perry KJ, Henry JQ (2015) CRISPR/Cas9-mediated genome modification in the mollusc, *Crepidula fornicata*. *Genesis (2000)*, **53**, 237–244.

Piatek A, Ali Z, Baazim H *et al.* (2014) RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnology Journal*, **13**, 578–589.

Piertney SB, Webster LMI (2010) Characterising functionally important and ecologically meaningful genetic diversity using a candidate gene approach. *Genetica*, **138**, 419–432.

Qi LS, Larson MH, Gilbert LA *et al.* (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173–1183.

Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, Gerard GF (2004) Mutation detection using Surveyor nuclease. *BioTechniques*, **36**, 702–707.

Radecke S, Radecke F, Cathomen T, Schwarz K (2010) Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, **18**, 743–753.

Ran FA, Hsu PD, Lin C-Y *et al.* (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, **154**, 1380–1389.

Rebeiz M, Williams TM (2011) Experimental approaches to evaluate the contributions of candidate cis-regulatory mutations to phenotypic evolution. *Methods in Molecular Biology (Clifton, NJ)*, **772**, 351–375.

Rebeiz M, Pool JE, Kassner VA, Aquadro CF, Carroll SB (2009) Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science (New York, NY)*, **326**, 1663–1667.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution*, **16**, 351–358.

Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T (2014) Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Scientific Reports*, **4**, 5400.

Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology*, **32**, 347–355.

Shin J, Chen J, Solnica-Krezel L (2014) Efficient homologous recombination-mediated genome engineering in zebrafish using TALE nucleases. *Development (Cambridge, United Kingdom)*, **141**, 3807–3818.

St Johnston D (2002) The art and design of genetic screens: *Drosophila melanogaster*. *Nature Reviews. Genetics*, **3**, 176–188.

Stapley J, Reger J, Feulner PGD *et al.* (2010) Adaptation genomics: the next generation. *Trends in Ecology & Evolution (Personal Edition)*, **25**, 705–712.

Storz JF, Wheat CW (2010) Integrating evolutionary and functional approaches to infer adaptation at specific loci. *Evolution*, **64**, 2489–2509.

Till BJ, Burtner C, Comai L, Henikoff S (2004) Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Research*, **32**, 2632–2641.

Torres R, Martin MC, Garcia A, Cigudosa JC, Ramirez JC, Rodriguez-Perales S (2014) Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. *Nature Communications*, **5**, 3964.

Tufts DM, Natarajan C, Revsbech IG *et al.* (2015) Epistasis constrains mutational pathways of hemoglobin adaptation in high-altitude pikas. *Molecular Biology and Evolution*, **32**, 287–298.

Waldman AS (2008) Ensuring the fidelity of recombination in mammalian chromosomes. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, **30**, 1163–1171.

Weston D, Patel B, Van Voorhis WC (1999) Virulence in Trypanosoma cruzi infection correlates with the expression of a distinct family of sialidase superfamily genes. *Molecular and Biochemical Parasitology*, **98**, 105–116.

Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, **482**, 331–338.

Wu S, Ying G, Wu Q, Capecchi MR (2008) A protocol for constructing gene targeting vectors: generating knockout mice for the cadherin family and beyond. *Nature Protocols*, **3**, 1056–1076.

Yang H, Wang H, Jaenisch R (2014) Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nature Protocols*, **9**, 1956–1968.

Zhang X, Wang H, Li M *et al.* (2014) Isolation of doublesex- and mab-3-related transcription factor 6 and its involvement in spermatogenesis in tilapia. *Biology of Reproduction*, **91**, 136.

All authors contributed to discussions about the content and organization of the review. J.M.B. helped curate the literature and was primarily responsible for writing the manuscript; E.C.O. helped curate the literature and wrote sections of the manuscript; L.M.M. helped to edit and revise the manuscript and prepared the figures.
