Available online at www.sciencedirect.com

ScienceDirect

Journal of Genetics and Genomics 41 (2014) 7−19

JGG

REVIEW

CRISPR/Cas9 and Genome Editing in *Drosophila*

Andrew R. Bassett*, Ji-Long Liu*

MRC Functional Genomics Unit, University of Oxford, Department of Physiology, Anatomy and Genetics, South Parks Road, Oxford OX1 3QX, United Kingdom

Received 25 November 2013; revised 10 December 2013; accepted 11 December 2013
Available online 18 December 2013

ABSTRACT

Recent advances in our ability to design DNA binding factors with specificity for desired sequences have resulted in a revolution in genetic engineering, enabling directed changes to the genome to be made relatively easily. Traditional techniques for generating genetic mutations in most organisms have relied on selection from large pools of randomly induced mutations for those of particular interest, or time-consuming gene targeting by homologous recombination. *Drosophila melanogaster* has always been at the forefront of genetic analysis, and application of these new genome editing techniques to this organism will revolutionise our approach to performing analysis of gene function in the future. We discuss the recent techniques that apply the CRISPR/Cas9 system to *Drosophila*, highlight potential uses for this technology and speculate upon the future of genome engineering in this model organism.

KEYWORDS: *Drosophila melanogaster*; CRISPR; Cas9; Genome engineering; Targeted mutagenesis

INTRODUCTION

The advent of genome sequencing and genome-wide technologies for study of gene expression, polymorphism and regulation has revolutionalised our ability to associate genes with particular cellular functions or disease states (McCarthy et al., 2008; Park, 2009). They have also allowed us to make predictions about the function of a large proportion of both coding and non-coding sequences (Roy et al., 2010; Bernstein et al., 2012). Although various techniques such as homologous gene targeting have allowed us to selectively mutagenise or alter gene function in a desired manner (Rong and Golic, 2000), the difficulty of applying these techniques on a large scale has restricted our ability to test hypotheses generated from such genome-wide analyses.

Genome editing systems

Genome editing technologies have been developed over the past decade that allow us to selectively mutagenise specific regions of the genome, and allow sophisticated and detailed mechanistic studies to be performed in a variety of organisms including *Drosophila* (Beumer et al., 2006, 2008; Liu et al., 2012). These technologies rely on specific DNA binding factors that can be used to target various functional domains to defined regions of the genome. Most experiments have used these reagents to generate a double strand break (DSB) in the DNA at the target site, that can then be repaired by non-homologous end joining (NHEJ) or homologous recombination (HR) (Bibikova et al., 2002). NHEJ is somewhat error-prone, and can result in the deletion or insertion of a few bases at the cut site, resulting in mutation of the DNA (Bibikova et al., 2002). HR normally results in precise repair

Abbreviations: CRISPR, clustered regularly interspaced short palindromic repeat; HRMA, high resolution melt analysis; PAM, protospacer adjacent motif; crRNA, CRISPR RNA; tracrRNA, trans-acting crRNA; Cas, CRISPR-associated; gRNA, guide RNA; sgRNA, single guide RNA.

* Corresponding authors. Tel: +44 1865 282 554, fax: +44 1865 282 849 (A.R. Bassett); Tel: +44 1865 285 833, fax: +44 1865 285 862 (J.-L. Liu).
E-mail addresses: andrew.bassett@dpag.ox.ac.uk (A.R. Bassett); jilong.liu@dpag.ox.ac.uk (J.-L. Liu).

http://dx.doi.org/10.1016/j.jgg.2013.12.004

from the sister chromatid, but if an excess of a desired homologous template is supplied, this may be used to introduce defined changes in the underlying DNA (Bibikova et al., 2003; Beumer et al., 2006, 2008).

The discovery that zinc finger DNA binding domains had a modular structure, each binding to three bases within the DNA provided the opportunity to select these proteins *in vitro* to bind to desired sequences within the genome (Choo and Klug, 1994a, 1994b). These reagents, when fused to a non-specific endonuclease domain, have been used successfully in various organisms, including flies, to induce DSBs at desired sites, resulting in targeted mutagenesis (Bibikova et al., 2002, 2003; Beumer et al., 2006, 2008). Direct injection into *Drosophila* embryos was shown to be able to introduce small insertions or deletions (indels). Upon coinjection of a homologous template DNA, this also allowed relatively efficient gene targeting (Beumer et al., 2008, 2013), which until this point was a time consuming and labour intensive process in *Drosophila* (Rong and Golic, 2000). However, the fact that the binding preference of each zinc finger depends on its context within the polypeptide (Del Rio et al., 1993; Bulyk et al., 2002) makes prediction of the binding sites of multimeric zinc fingers difficult, hindering a more general application of this technique.

A second class of highly modular DNA binding proteins termed transcription activator like effectors (TALEs) were discovered in the plant pathogen *Xanthomonas* (Bonas et al., 1989; Boch et al., 2009; Moscou and Bogdanove, 2009). They are made of 34 amino acid repeating units, each binding to a single base of DNA in a highly predictable manner, dependent on the identity of two amino acids (the repeat variable diresidue or RVD) within each unit (Boch et al., 2009; Moscou and Bogdanove, 2009). DNA binding of each unit was essentially independent of its context, and the number of monomer units can be adjusted to determine the length of the DNA binding site (Cermak et al., 2011; Li et al., 2011). This provides exquisite specificity in DNA binding, but assembly of these repetitive multimers requires complex systems of cloning (Cermak et al., 2011; Reyon et al., 2012; Sanjana et al., 2012), and the relatively large size of the resulting polypeptides makes their manipulation and expression somewhat problematic in some contexts.

CRISPR/Cas9

The clustered regularly interspaced short palindromic repeat (CRISPR/Cas9) system acts as a bacterial defense system against invading viruses and plasmids in many different bacterial species (Fig. 1 and Table S1) (Ishino et al., 1987; Jansen et al., 2002; Barrangou et al., 2007; Garneau et al., 2010). The best studied system is that from *Streptococcus pyogenes*. Here, the Cas9 endonuclease is targeted to sequences from the invading pathogen by a crRNA (CRISPR RNA), that provides specificity to the endonuclease by base pairing with a 20 nt complimentary sequence within the DNA (Brouns et al., 2008; Gasiunas et al., 2012; Jinek et al., 2012). Endogenously, a further component, known as the tracrRNA (trans-acting crRNA) forms a complex with the crRNA and targets its incorporation into the Cas9 complex. Recently, this system has been shown to work in many other organisms, including mammalian (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013b; Wang et al., 2013), insect (Bassett et al., 2013, 2014; Gratz et al., 2013; Kondo and Ueda, 2013; Ren et al., 2013; Sebo et al., 2013; Yu et al., 2013), plant (Shan et al., 2013) and fungal (DiCarlo et al., 2013) cells. Fusion of the crRNA and tracrRNA into a ~100 nt synthetic single guide or chimeric RNA (sgRNA or chiRNA) has further simplified this system, which then only requires two components to be expressed (Dahlem et al., 2012; Jinek et al., 2012; Cong et al., 2013; Mali et al., 2013b) (Fig. 2). The specificity is determined by a 20 nt sequence at the 5′ end of the sgRNA, which can be altered to match any desired sequence in the DNA. The only limitation upon this targeting is that the 20 nt guide sequence has to be followed by a protospacer adjacent motif (PAM) of NGG in the DNA in order for efficient cleavage to occur (Fig. 2). This sequence should occur on average every 8 bases in the DNA, but recent reports have suggested that this requirement may be relaxed to include NAG sequences (Mali et al., 2013a), increasing the number of potential target sites still further. CRISPR systems from other bacterial species have different PAM requirements, for example CWT and GAA (Mojica et al., 2009; Esvelt et al., 2013) and this suggests that it will be possible to engineer Cas proteins to bind to essentially any sequence in the future.

DROSOPHILA CRISPR SYSTEMS

Several groups have used the CRISPR/Cas9 system to induce targeted mutations in *Drosophila* (Bassett et al., 2013; Gratz et al., 2013; Kondo and Ueda, 2013; Ren et al., 2013; Sebo et al., 2013; Yu et al., 2013), but differ in their approach to supplying the Cas9 protein and sgRNA components of the system (Tables 1 and 2).

The first description of mutagenesis with CRISPR/Cas9 involved coinjection of two plasmids into syncytial blastoderm stage *Drosophila* embryos (Gratz et al., 2013). One plasmid expresses the *Cas9* gene under the *Hsp70* promoter, and the second produces the sgRNA, driven by a *pol III* promoter from the *U6* gene. This was tested at the *yellow* gene, and resulted in mutagenesis of the gene that was capable of being transmitted to subsequent generations. The efficiency of mutagenesis due to inefficient NHEJ was fairly low, with 5.9% of the injected flies giving rise to at least one mutant offspring (Table 1) (Gratz et al., 2013). However, the authors further demonstrated that if two sgRNAs are supplied, targeting either end of the *yellow* gene, this can result in deletion of the intervening sequence, and that integration of short sequences at the cleavage site is possible by coinjection with a short single stranded oligonucleotide donor sequence (Gratz et al., 2013).

A second technique that has been applied by two groups independently involves coinjection of *in vitro* transcribed *Cas9* mRNA and sgRNA into early stage embryos, and achieves much higher mutagenesis rates due to inefficient NHEJ (Bassett et al., 2013; Yu et al., 2013). Bassett et al. showed that up to 88% of injected flies gave rise to mosaic expression of the *yellow* gene implying that this technique is highly efficient.

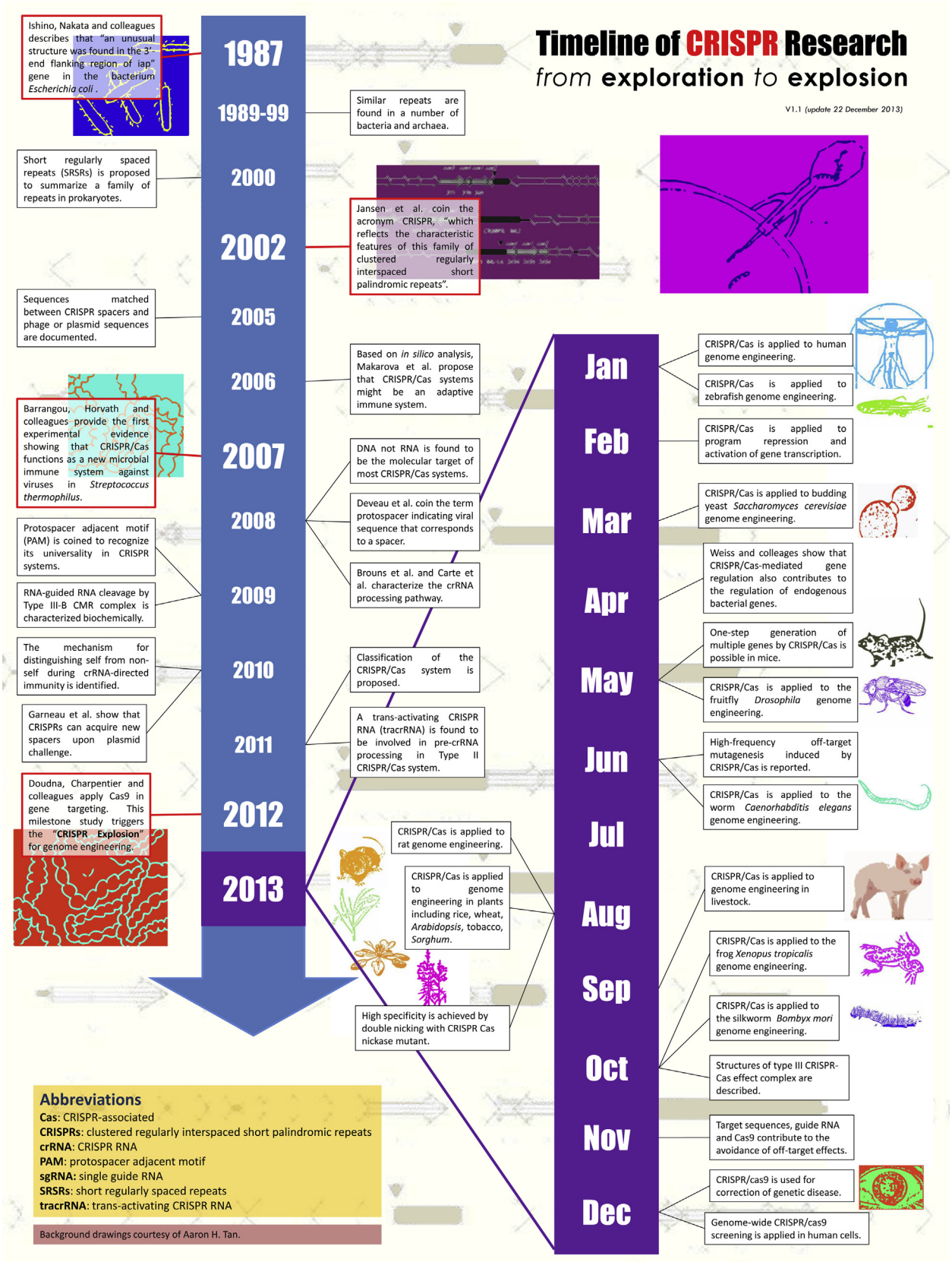

Fig. 1. Timeline of CRISPR research − from exploration to explosion.
Four milestone studies are highlighted in red boxes. See Table S1 for a more detailed timeline with references. Background drawings courtesy of Aaron H. Tan.

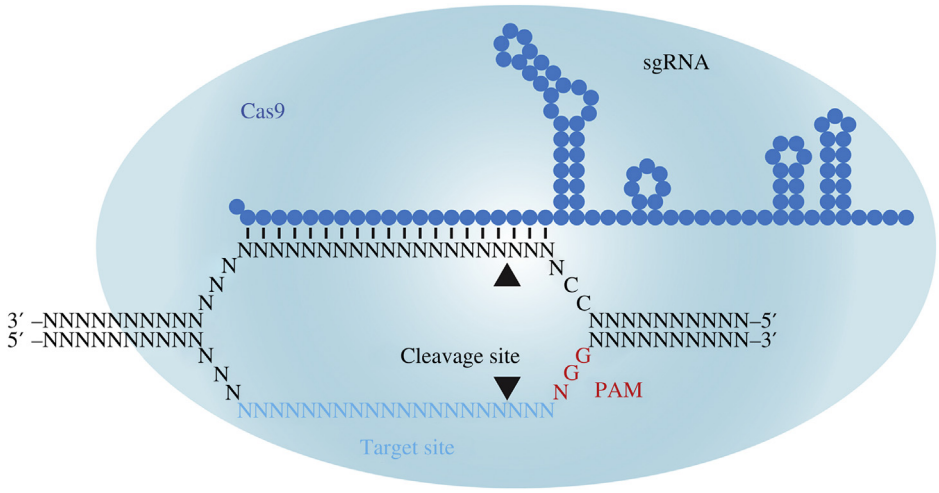

Fig. 2. Two-component CRISPR/Cas9 system for genome editing.
The Cas9 protein (blue oval) is recruited to a target site in the DNA by a 20 nt complimentary sequence in the synthetic single guide RNA (sgRNA, dark blue). Cleavage requires a protospacer adjacent motif (PAM (NGG), red) in the DNA, which does not appear in the sgRNA. A double strand break is made 3 nt from the PAM sequence on both strands of the DNA (cleavage site, black triangles).

Table 1
Comparison of CRISPR genome engineering studies in *Drosophila*

Reference	Gratz et al., 2013	Bassett et al., 2013	Yu et al., 2013	Kondo and Ueda, 2013	Sebo et al., 2013	Ren et al., 2013
Cas9 promoter	*hsp70*	T7	Sp6	*nos*	*vasa*	*nos*
Delivery	DNA injection	mRNA injection	mRNA injection	Transgenic	Transgenic	Transgenic
sgRNA promoter	*U6*	T7	T7	*U6*	*U6*	*U6a*, *U6b*, *nos*-mini
Delivery	DNA injection	sgRNA injection	sgRNA injection	Transgenic	DNA injection	DNA injection
Gene targeted	*yellow*	*yellow*, *white*	*yellow*, *K81*, *CG3708*, *CG9652*, *kl-3*, *light*, *RpL15*	*white*, neuropeptide genes (*Ast*, *capa*, *Ccap*, *Crz*, *Eh*, *Mip*, *npf*), *mir-219*, *mir-315*	*EGFP*, *mRFP*	*white*
Mutant detection	Body colour, surveyor assay, DNA sequencing	Body colour, eye colour, DNA sequencing, HRMA	Body colour, DNA sequencing	Eye colour, T7 endonuclease, DNA sequencing	Fluorescence in the eye	Eye colour, HRMA, DNA sequencing
G_0 adult fertility (%)[a]	nd	37.5−73.1	21.6−94.4	N/A	32−55	29.4−92.9
Mosaic G_0 (%)[b]	6−66	4−88	35.7−80	N/A	0[g]	0[g]
Germline mutants (among fertile flies) (%)[c]	5.9−20.7	0−79	35.7−100	0−100	35−71	0−100
Germline mutants (of injected flies) (%)[d]	nd	0−58	nd	N/A	N/A	0−81.2
G_1 mutant rate per fly (%)[e]	nd	0−88.5	N/A	N/A	0−100	0−100
G_1 mutant overall (%)[f]	0.25−1.37	0−34.5	2.1−98.9	0−99.4	7.7−24.7	0−74.2
Off-target detection method	nd	HRMA	nd	nd	nd	HRMA, DNA sequencing
Off-target detected	nd	No	nd	nd	nd	No
Homologous recombination	Oligonucleotide	nd	nd	nd	nd	nd
Long deletions	Yes	No	No	Yes	No	Yes

[a] Number of fertile flies as a proportion of the surviving injected flies; [b] Percentage of flies that exhibit mosaic expression in the injected generation, either visibly in males or detected using HRMA (high resolution melt analysis); [c] Proportion of fertile flies giving rise to at least one mutant offspring; [d] Proportion of injected flies giving rise to at least one mutant offspring; [e] For individual flies, the range of the percentage of mutant offspring; [f] Total number of mutant G_1 offspring as a percentage of the total offspring; [g] No mosaic expression would be expected, due to germline expression of Cas9; N/A, not applicable to this technique; nd, not determined in this study.

Table 2
Web resources for CRISPR/Cas study

Name	Link	Brief description	Reference
***Drosophila* CRISPR web resources**			
OxfCRISPR (Liu Lab)	http://www.oxfcrispr.org/	Oxford Fly CRISPR Resources	Bassett et al., 2013
CRISPRflydesign (Bullock Lab)	http://www.crisprflydesign.org/	Offers Cas9 transgenic stocks	n/a
DRSC CRISPR finder (Perrimon Lab)	http://www.flyrnai.org/crispr/	A web tool to identify CRISPRs for fly study	Ren et al., 2013
FlyCas9 (Ueda Lab)	http://www.shigen.nig.ac.jp/fly/nigfly/cas9/index.jsp	Provides reagents, protocols and online tools for genome engineering by the designer nuclease Cas9 in *Drosophila*	Kondo and Ueda, 2013
flyCRISPR (O'Connor-Giles Lab, Wildonger Lab and Harrison Lab)	http://flycrispr.molbio.wisc.edu/	Fly CRISPR resources	Gratz et al., 2013
flyCRISPR discussion group	https://groups.google.com/forum/#!forum/flycrispr-discussion-group	A forum for sharing developments, insights, ideas and asking questions related to fly CRISPRs	Gratz et al., 2013
Fly target sites (Zhang Lab)	http://www.genome-engineering.org/crispr/?page_id=41	For application of Cas9 for site-specific genome editing in eukaryotic cells and organisms	Hsu et al., 2013
General CRISPR resources			
Addgene CRISPR plasmids	http://www.addgene.org/CRISPR/	A collection of CRISPR plasmids and reagents	n/a
Crass: The CRISPR Assembler	http://ctskennerton.github.io/crass/	A program that searches through raw metagenomic reads for CRISPRs	Skennerton et al., 2013
CRISPI	http://crispi.genouest.org/	A web interface with graphical tools and functions allows users to find CRISPR in personal sequences.	Rousseau et al., 2009
CRISPR Discussion Forum	https://groups.google.com/forum/#!forum/crispr	A forum to discuss Genome Engineering using CRISPR/Cas systems	Cong et al., 2013
CRISPRmap	http://rna.informatik.uni-freiburg.de/CRISPRmap	Web server provides an automated assignment of newly sequenced CRISPRs to standard classification system	Lange et al., 2013
CRISPRs web server	http://crispr.u-psud.fr/	A gateway to publicly accessible CRISPRs database and software including CRISPRFinder, CRISPRdb and CRISPRcompar	Grissa et al., 2007a, 2007b, 2008
CRISPRTarget	http://bioanalysis.otago.ac.nz/CRISPRTarget	Predicts the most likely targets of CRISPR RNAs	Biswas et al., 2013
E-CRISP	http://www.e-crisp.org/	A software tool to design and evaluate CRISPR target sites	n/a
Goldstein Lab CRISPR	http://wormcas9hr.weebly.com/	A genome engineering resource for the *Caenorhabditis elegans* research community	Dickinson et al., 2013
Joung Lab CRISPR	http://www.crispr-cas.org/	A genome engineering resource for zebrafish research community	Hwang et al., 2013
Zhang Lab Genome Engineering	http://www.genome-engineering.org/	CRISPR genome engineering resources website	Cong et al., 2013
ZiFiT target design tool	http://zifit.partners.org/ZiFiT/	Identifies potential target sites in DNA sequences	Sander et al., 2007, 2010

n/a, not available.

A second study by Yu et al. showed a similar efficiency (80%) at the *yellow* gene, but also showed successful mutagenesis at six other target loci spread throughout the genome, demonstrating the general applicability of this approach. The difference in efficiency between plasmid and mRNA injection techniques may be explained by the expression levels of the Cas9 protein and sgRNA, or by the timing of expression relative to the specification of germ cells in the embryo.

A third system has also been developed whereby two transgenic flies are produced, one expressing Cas9 in the germline under the *nanos* promoter, and a second with ubiquitous expression of the sgRNA again driven by the *U6* promoter (Kondo and Ueda, 2013). When these two flies are crossed together, highly efficient mutagenesis can be achieved, giving rise to up to more than 90% of flies with at least one mutant offspring (Table 1), and allowing longer deletions of up to 1.6 kb to be made efficiently by coexpression of two sgRNAs. Although efficient, this requires the time consuming step of producing a new transgenic fly for each sgRNA required, and removal of the Cas9 and sgRNA transgenes after mutant generation. However, this technique will have advantages in certain applications, since it is more reproducible than the techniques involving embryo injection.

The final technique uses injection of plasmids encoding the sgRNA into transgenic lines in which Cas9 is expressed specifically in the germline under the *vasa* (Sebo et al., 2013) or *nanos* (Ren et al., 2013) promoters. These techniques avoid potentially problematic somatic mutagenesis by limiting Cas9 expression to the germline cells. Sebo et al. (2013) demonstrated high rates of mutagenesis in the G_1 offspring derived from flies injected with plasmids encoding sgRNAs, but a significant proportion of the injected flies were infertile. By using the *nanos* promoter to drive Cas9 expression, Ren et al. (2013) achieved higher rates of fertility, and generated high rates of mutagenesis in G_1 offspring (Table 1).

In addition to the injection of vectors expressing guide RNAs (gRNAs), the production of fly lines that express Cas9 either ubiquitously or in the germline will also allow direct injection of *in vitro* transcribed sgRNA into these embryos. However, the relative efficiency of this technique has not been established.

Recently, expression vectors for Cas9 expression in *Drosophila* cell lines have also been described. The *Actin5c* and *U6* promoters were use to drive expression of the Cas9 and sgRNA components, respectively. This results in highly efficient mutagenesis in more than 80% cells due to the indels generated by inefficient NHEJ. The authors also demonstrated that homologous integration is possible using short oligonucleotide donors to insert small sequences, or longer homology arms to insert a 1.8 kb cassette at up to 4% efficiency (Bassett et al., 2014).

USES OF CAS9 INDUCED DSBS IN GENOME ENGINEERING

The majority of applications of CRISPR/Cas9 in genome engineering use its ability to introduce DSBs at specific sites

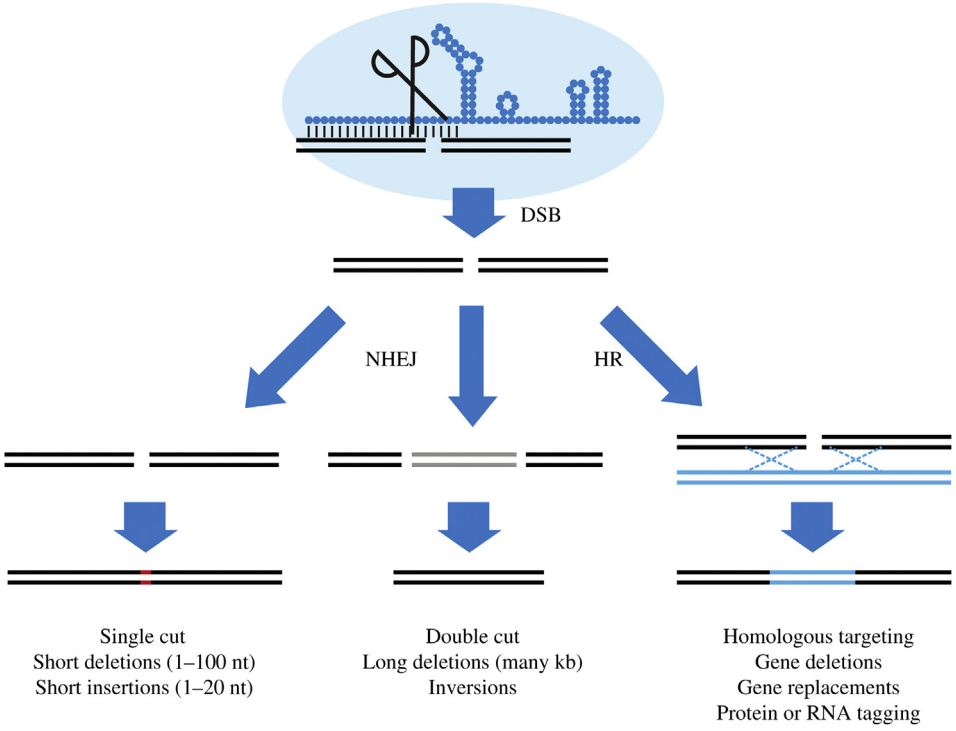

Fig. 3. Double strand break repair can be used to target defined genomic changes.
The double strand break (DSB) induced by the Cas9/sgRNA complex can be repaired by non-homologous end joining (NHEJ) or homologous recombination (HR). This can result in small insertions or deletions at the target site (left), deletions of larger genomic regions when two cuts are made (middle) or homologous repair with a desired template (right). This can be used to alter the genome in a variety of different ways (bottom).

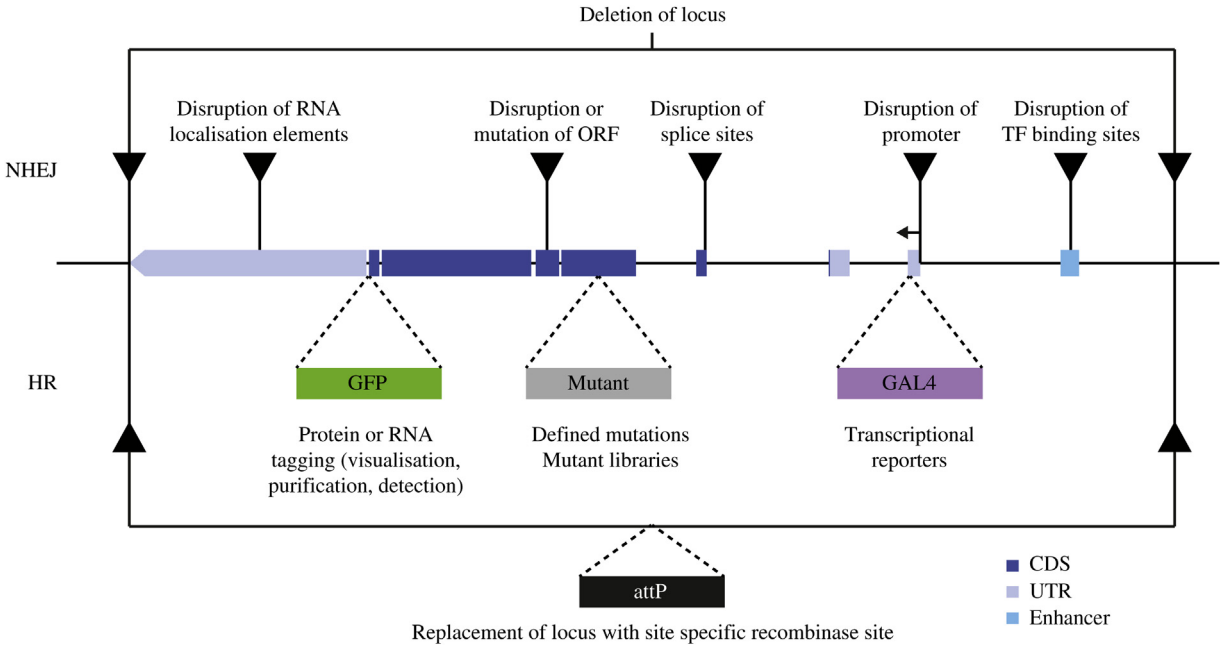

Fig. 4. Uses of genome editing within protein coding genes.
NHEJ repair (top) of Cas9-induced DSBs can be used to remove functional elements or disrupt genes in a variety of different ways. HR repair (bottom) can be used to insert or replace sequences present within the gene for a variety of different uses. Exons are indicated as boxes, with coding sequence (CDS) in dark blue and untranslated region (UTR) in grey. Enhancers are also indicated in light blue.

within the genome. The DSBs can be repaired by either NHEJ or HR, and both repair mechanisms can be used to generate mutations and manipulate the genome in a defined manner (Figs. 3 and 4).

NHEJ

Mistakes during the NHEJ repair process often result in small indels, and these can be useful in mutagenesis by shifting reading frame of protein coding genes, resulting in null mutations or truncations of the protein sequence. It is also possible to use the indels generated to disrupt other small functional sites within the genome such as transcription factor binding sites, miRNA (microRNA) genes, splice donor or acceptor sites and critical amino acids within protein coding sequence (Fig. 4). These kinds of mutation can be used to define regulatory regions, investigate miRNA function, or force different splicing patterns to occur as well as generating mutations within protein coding genes.

Pairs of nucleases can also be used to trigger deletion or inversion of intervening sequences (Gupta et al., 2013; Xiao et al., 2013) (Fig. 3). This can be useful to make defined deficiencies or remove larger functional regions, such as protein domains, non-coding RNA genes or larger regulatory elements from the genome, or investigate the effects of larger chromosomal rearrangements upon the function of the organism.

Uses of NHEJ-mediated mutation

The development of CRISPR/Cas9 for genome engineering in *Drosophila* can be used to generate mutations in protein coding genes, especially those that have been refractory to current technologies such as P-element mutagenesis. Despite the large number of more than 30,000 transposon insertions and other mutants generated in this organism, only 40%−50% of annotated protein coding genes contain an insertion, and although many of these have orthologs in other species, their function remains unknown (Bellen et al., 2004, 2011).

Only 20% of the *Drosophila* genome codes for proteins (Lin et al., 2007), yet around 50% of the sequence displays evolutionary conservation (Meader et al., 2010), suggesting that it may be functional. Targeted mutagenesis can also be applied to study of the regulatory or other functional elements present within this sequence, such as transcription factor binding sites, non-coding RNAs and other non-coding elements.

The ability to inject into essentially any genetic background (Bassett et al., 2013; Yu et al., 2013) will also allow the compounding of multiple mutations, especially those that are tightly linked genetically. This will enable studies of redundancy amongst members of a protein family, or epistasis or genetic interactions between tightly linked genes. We have also shown that coinjection of two sgRNAs can result in simultaneous mutation of two target genes, making such analyses even easier to achieve (Fig. 5). Being able to create mutants in a defined genetic background will also benefit studies of more subtle phenotypes such as behaviour, that are highly dependent on the background used, relieving the need for time consuming backcrossing. The fact that only two components need to be supplied, and that RNA injection techniques can be applied to multiple *Drosophila* species will also make it useful for evolutionary and developmental studies

to recapitulate or validate the functionality of genomic changes between different species.

The high efficiency of homozygous mutant generation by CRISPR/Cas9 systems will also make it possible to perform mosaic analysis either by embryo injection of RNA (Bassett et al., 2013; Ren et al., 2013; Sebo et al., 2013; Yu et al., 2013) or crosses between Cas9 and sgRNA transgenes (Kondo and Ueda, 2013). Homozygous mutant patches can be marked with an antibody to the protein of interest, and used to study phenotypes in an otherwise wild type background. It may also be possible to generate tissue, cell and developmental stage-specific knockouts by expression of the Cas9 protein in specific patterns using the UAS−GAL4 system, and subsequent crossing these flies to flies expressing the sgRNA of choice. This kind of system may be useful to investigate essential genes whose knockout results in early lethality.

HR

Creation of a DSB increases the rate of homologous repair at that site by several orders of magnitude (Gloor et al., 1991; Choulika et al., 1995; Smih et al., 1995), and this enables gene targeting to produce defined genetic changes much more rapidly and quickly than with classical techniques (Rong and Golic, 2000) (Fig. 3). This relies on supplying a large excess of a homologous repair template with the desired changes (Beumer et al., 2008, 2013). The donor DNA (ssDNA) can take two forms: single stranded DNA (ssDNA) oligonucleotides synthesised up to 200 nt in length and used to integrate short sequences, or longer double stranded DNA (dsDNA) constructs containing hundreds to thousands of nucleotides of homologous sequence on either side of the DSB site (Beumer et al., 2013). The latter are capable of integrating longer sequences at higher efficiency. The majority of repair events in fly cells occur by NHEJ, but repair can be biased towards HR by using a mutation in an essential component of the NHEJ pathway, DNA ligase 4 (*lig4*) (Beumer et al., 2008, 2013; Bassett et al., 2014). This reduces the number of non-homologous integrants significantly, greatly improving the efficiency of homologous gene targeting.

Integration of short sequences with ssDNA oligonucleotides

Integration of ssDNA oligonucleotides allows sequences such as epitope tags to be integrated into protein coding genes for detection of endogenous proteins, and defined mutations to be generated (Fig. 4). This can be useful to generate disease models, investigate the function of specific amino acids in proteins or recapitulate polymorphisms in coding or non-coding regions for studies of evolution.

Site-specific recombinase sites such as the widely used attP sites (Groth et al., 2004) can also be integrated to allow subsequent modifications at the same position to be easily achieved. This technique can also be combined with two sgRNAs to remove a genomic region and replace it with a site-specific recombinase site (Gratz et al., 2013). This allows subsequent

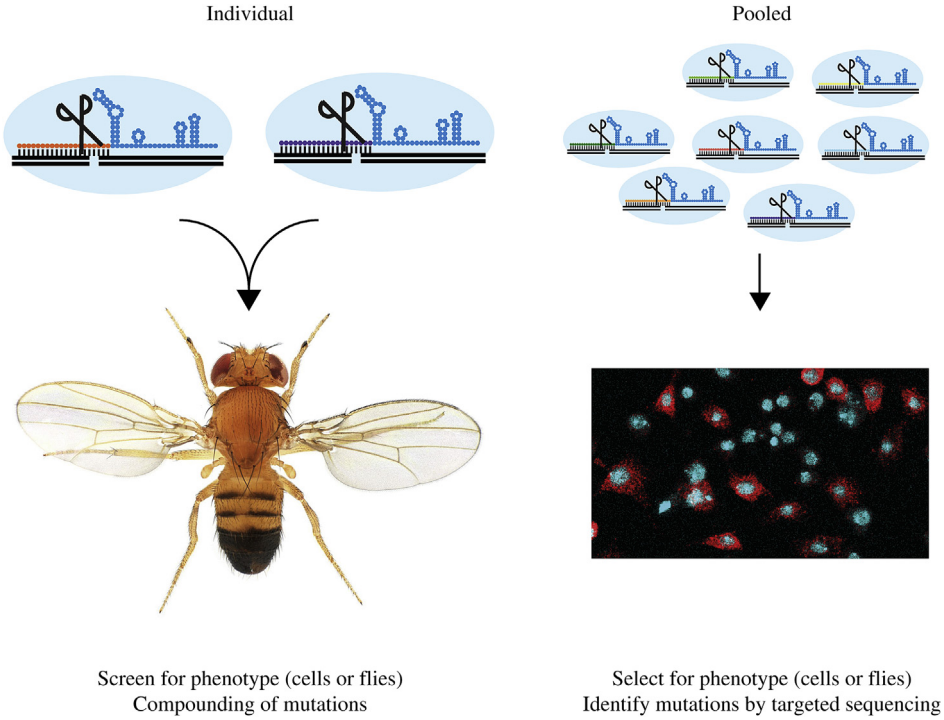

Fig. 5. Genetic screening with CRISPR/Cas9 can be performed on individual cases or in a pooled manner.
CRISPR/Cas9 nucleases can be used individually in a one-by-one manner (left) to disrupt the function of a gene and measure phenotypes in flies or cells. They can also be combined to disrupt more than one gene simultaneously as indicated by the mosaic *yellow* and *white* expression in the indicated fly. The nucleases can also be pooled (right) and transformed into cells or injected into flies. After selection or screening for a phenotype of interest, mutations can be identified by sequencing of the genomic regions targeted by the sgRNAs.

replacement of the deleted region with modified versions to perform detailed investigations of its function. Other recombinase sites such as flippase recognition target (FRT) sites (Golic and Lindquist, 1989) could be introduced flanking the region of interest to allow tissue specific or clonal deletion of genomic regions (Fig. 4).

Integration of other sequences with long dsDNA donors

Longer dsDNA homology arms can also be used in the targeting vector to integrate longer sequences at higher efficiency than short oligonucleotides (Beumer et al., 2013). This allows use of a marker gene to allow simple selection for homologous integrants, which is particularly advantageous as the efficiency of this process is still relatively low. Studies in both flies (Beumer et al., 2013) and S2 cells (Bassett et al., 2014) have suggested that 1 kb of homology at either side of the integration site directs efficient homologous integration, which makes construction of the targeting constructs simple, and detection of integration possible by PCR. As well as integration of a positive marker for integration, homologous targeting in *Drosophila* has been optimised by incorporation of negative selectable markers such as UAS-*rpr* outside of the homology arm to enable selection against non-homologous integrations (Huang et al., 2008; Baena-Lopez et al., 2013). These techniques could be applied to improve efficiency still further.

Larger homology arms enable integration of longer features, for example fluorescent proteins such as GFP to visualise expression of proteins, GAL4 as a transcriptional reporter to analyse expression patterns of transcripts or affinity tags for efficient protein purification (Fig. 4). Tagging at the endogenous locus would ensure that such proteins would be subject to the same transcriptional and post-transcriptional controls as the endogenous gene, and one would be sure of their functionality and expression level.

TARGETED GENETIC SCREENING

The simplicity at which sgRNAs can be made also offers the opportunity to create genome-wide libraries targeting every protein coding gene. This would provide a resource of specific mutagens that target only the regions of interest, rather than the random mutagens currently used such as ethyl methanesulfonate (EMS), X-rays or P-elements. This provides a revolutionary approach to genetic screening, both in cell culture and in the context of a whole organism.

Due to the ease at which the mutations can be identified, this can be performed on a pooled collection of sgRNAs as well as using a one-by-one targeted approach (Fig. 5). This offers an alternative to the genome-wide RNAi libraries currently available in cells and that are being generated in flies (Flockhart et al., 2006; Flockhart et al., 2012). These are only able to provide a partial and post-transcriptional loss of function rather than genetic knockouts.

Cell-based systems can be used to screen for mutations involved in selectable cellular phenotypes, but perhaps one of the most exciting possibilities would apply genome wide libraries of sgRNAs in the context of the whole organism (Shalem et al., 2014; Wang et al., 2014). This would enable screens for phenotypes resulting from gene knockout or to perform genetic modifier screening far more efficiently. This would allow rigorous genetic screens that knock out the function of every protein coding gene in a defined manner. It would also be possible to perform semi-targeted mutagenesis by selecting subsets of sgRNAs based on expression patterns in the tissue of interest, subcellular distributions or involvement in particular biological processes.

EXTENDED USES OF THE CAS9 SYSTEM

Recently, there have been many developments of the Cas9 system, other than the introduction of DSBs. These use a mutant Cas9 protein that is catalytically inactive and unable to cleave the DNA, resulting in a sequence specific DNA binding factor that can be targeted to virtually any genomic region by coexpression with a sgRNA (Qi et al., 2013).

The simplest use of this system is to sterically interfere with the process of transcription by targeting the Cas9−sgRNA complex to the middle of a transcribed region, known as CRISPRi (Qi et al., 2013) (Fig. 6). This has been demonstrated to interfere with the progress of RNA polymerase and reduce transcription levels in both bacterial and mammalian cells, and could be applied to any other system.

Furthermore, the Cas9 protein can be fused to other proteins of interest in order to target them to specific sites within the DNA. These could be fluorescent proteins to tag specific DNA sequences for live imaging, or affinity tags for purification of specific regions of chromatin. Transcriptional activation and repression has been demonstrated by fusion of Cas9 to multiple copies of the strong activation domain (VP64) (Cheng et al., 2013; Farzadfard et al., 2013; Mali et al., 2013a; Perez-Pinera et al., 2013) or Krüppel-associated box (KRAB) repression domain (Farzadfard et al., 2013; Gilbert et al., 2013) in human cells, and this could be used to regulate gene expression without introduction of exogenous sequences. It would also be possible to recruit chromatin modifying complexes, enabling addition or removal of post-transcriptional modifications of histones and DNA or ATP-dependent chromatin remodelling enzymes to alter nucleosome positions or to incorporate specific histone variants. This will enable study of directed changes to epigenetic states to investigate their roles *in vivo*.

Dimerisation domains could also be fused to Cas9, and pairs of Cas9 proteins targeted to discrete regions of the genome could force looping of the intervening DNA. This can be used to reorganise the topology of the underlying chromatin and study its effects on gene expression.

It has also been demonstrated that the sgRNA can be extended at its 3′ end, providing an opportunity to use the CRISPR/Cas9 system to recruit RNA molecules to sites within the DNA. Recent studies have postulated that some non-coding RNAs can recruit protein complexes to specific regions of the genome (Khalil et al., 2009; Zhao et al., 2010;

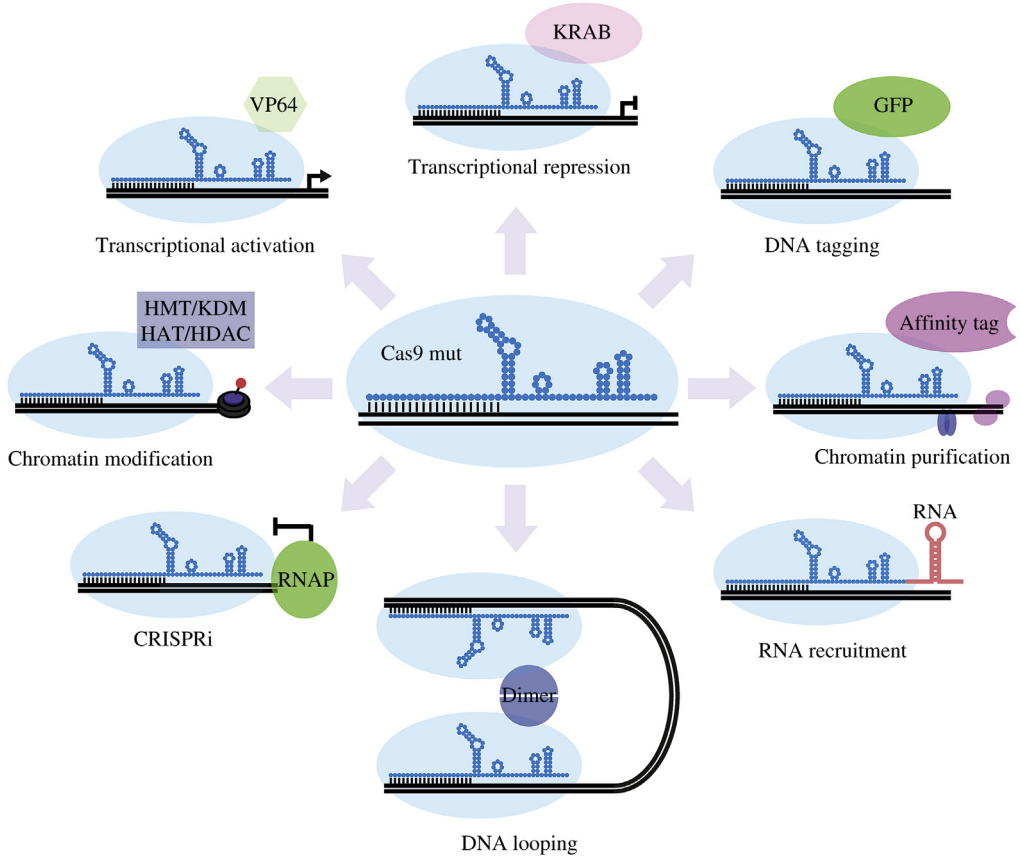

Fig. 6. Alternative uses of nuclease-deficient Cas9 protein.
Nuclease deficient Cas9 protein (Cas9 mut) can be used for multiple other functions including regulation of transcription by interference with RNA polymerase (CRISPRi) or fusion to transcriptional activation (VP64) or transcriptional repression (KRAB) domains. Other functions include DNA tagging with GFP, chromatin purification using an affinity tag, RNA recruitment by fusion to the sgRNA, or altering DNA topology by using dimerisation domains (DNA looping). Chromatin modification with histone methyltransferase (HMT), acetyl-transferase (HAT), demethylase (KDM) or deacetylase (HDAC) can also be targeted by fusion to the appropriate domains.

Brockdorff, 2013; Davidovich et al., 2013), and this system can be used to test the effect of artificial recruitment of RNAs to specific points within the genome to test such hypotheses. Given the exquisite specificity of Cas protein binding to the structure present within the gRNA, this may be fused to other RNAs as an affinity tag, and used to visualise or purify a RNA by virtue of its binding to the Cas protein (Lee et al., 2013).

LIMITATIONS OF THE CRISPR/CAS9 SYSTEM

A limitation of the CRISPR/Cas9 system is its relatively low targeting specificity, which is determined by a 20 nt recognition site and the requirement for the neighbouring PAM sequence (NGG) (Wei et al., 2013). This is further confounded by the fact that several mismatches within the target sequence can be tolerated, whilst still directing efficient cleavage (Fu et al., 2013; Mali et al., 2013a). However, due to its relatively small genome size, this is less of a problem in *Drosophila* than in other systems, and sequences can be carefully chosen that minimise such off-target effects. Several web-based algorithms to perform such analyses have recently been described (Ren et al., 2013) (Table 2). The short generation time of *Drosophila* also allows such off-target mutations to be removed relatively simply by backcrossing if specific target sites are not available. It would also be possible to use the "double-nick" approach to improve specificity. This uses a mutated Cas9 protein that is only able to make single strand nicks in the DNA. By judicious use of two sgRNAs that target nearby sequences, this doubles the sequence specificity, and additionally requires correct spacing between the two target sequences (Mali et al., 2013a; Ran et al., 2013).

Another issue with sgRNA design is that the efficiency of cleavage varies considerably at different target sites. This could be due to many reasons such as secondary structures within the sgRNA, thermodynamic stability of the sgRNA-DNA duplex or accessibility of the target sequence within the context of chromatin. Rigorous studies of such effects have not yet been performed, and it is therefore important to design multiple sgRNAs for each desired target to maximise chances of successful mutagenesis.

Recent observations have also suggested that over-expression of Cas9 alone with the actin−GAL4 driver can result in toxicity. This suggests that even in the absence of a sgRNA, there may be a degree of non-specific off-target mutagenesis, which should be borne in mind when analysing Cas9-induced mutations. As the technique becomes more

widely adopted, all of these problems will be better understood, allowing us to minimise their effects. Forums for distribution of such information are now available (Table 2), and will no doubt enhance our ability to apply these systems more generally.

CONCLUSION

The ability to use CRISPR/Cas9 systems for genome engineering in *Drosophila* and its cell lines will revolutionise genetic analysis by providing a targeted mutagen, which is ideally suited to a relatively small genome size. The simplicity at which it can be reprogrammed to target different sites will also enable large-scale application of this technology genome-wide, to allow more powerful reverse genetic analyses, and allow us to more easily study combinations of mutations.

A simpler method of homologous targeting will also enable wider application of defined modifications of the genome to study the importance of genomic regions. This can be used to make more appropriate tools for visualisation, detection and purification of proteins. Importantly, this will allow detailed analysis of gene function that would not be visible in simple genetic knockouts, and permit studies of the vast expanse of non-coding sequence.

The plethora of other techniques that use the sequence specific binding ability of Cas proteins to recruit other protein domains or RNA moieties offers a wealth of resources to manipulate transcriptional, epigenetic or topological features of chromatin. This will enable more delicate manipulations of many cellular functions, and the array of techniques available will no doubt be expanded in the future.

Genome editing will change the way in which we think about and perform genetic and genomic analysis. Combined with the power of genetic screens to identify phenotypes or modify existing ones, and the powerful developmental genetic tools and reagents already available, it will keep the humble fruit fly to the forefront of genetics for many years to come.

ACKNOWLEDGEMENTS

Further information concerning experimental methods and links to discussion groups and other information are provided at the OXfCRISPR website (http://oxfcrispr.org). We would like to thank Dr. Nicolas Gompel for the high resolution images of *Drosophila* in Fig. 5, Prof. Chris Ponting for his support, and Dr. Charlotte Tibbit and Dr. Sarah Cooper for critical reading of the manuscript. The CRISPR/Cas9 projects were supported by the UK Medical Research Council and the European Research Council (DARCGENs, No. 249869).

SUPPLEMENTARY DATA

Table S1. Timeline of CRISPR research related to Fig. 1.

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jgg.2013.12.004.

REFERENCES

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315, 1709−1712.

Baena-Lopez, L.A., Alexandre, C., Mitchell, A., Pasakarnis, L., Vincent, J.P., 2013. Accelerated homologous recombination and subsequent genome modification in *Drosophila*. Development 140, 4818−4825.

Bassett, A.R., Tibbit, C., Ponting, C.P., Liu, J.L., 2013. Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. Cell Rep. 4, 220−228.

Bassett, A.R., Tibbit, C., Ponting, C.P., Liu, J.L., 2014. Mutagenesis and homologous recombination in *Drosophila* cell lines using CRISPR/Cas9. Biol. Open 3, 42−49.

Bellen, H.J., Levis, R.W., He, Y., Carlson, J.W., Evans-Holm, M., Bae, E., Kim, J., Metaxakis, A., Savakis, C., Schulze, K.L., Hoskins, R.A., Spradling, A.C., 2011. The *Drosophila* gene disruption project: progress using transposons with distinctive site specificities. Genetics 188, 731−743.

Bellen, H.J., Levis, R.W., Liao, G., He, Y., Carlson, J.W., Tsang, G., Evans-Holm, M., Hiesinger, P.R., Schulze, K.L., Rubin, G.M., Hoskins, R.A., Spradling, A.C., 2004. The *BDGP* gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes. Genetics 167, 761−781.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., Snyder, M., 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57−74.

Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J.K., Carroll, D., 2006. Efficient gene targeting in *Drosophila* with zinc-finger nucleases. Genetics 172, 2391−2403.

Beumer, K.J., Trautman, J.K., Bozas, A., Liu, J.L., Rutter, J., Gall, J.G., Carroll, D., 2008. Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. Proc. Natl. Acad. Sci. USA 105, 19821−19826.

Beumer, K.J., Trautman, J.K., Mukherjee, K., Carroll, D., 2013. Donor DNA utilization during gene targeting with zinc-finger nucleases. G3 (Bethesda) 3, 657−664.

Bibikova, M., Beumer, K., Trautman, J.K., Carroll, D., 2003. Enhancing gene targeting with designed zinc finger nucleases. Science 300, 764.

Bibikova, M., Golic, M., Golic, K.G., Carroll, D., 2002. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. Genetics 161, 1169−1175.

Biswas, A., Gagnon, J.N., Brouns, S.J., Fineran, P.C., Brown, C.M., 2013. CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. RNA Biol. 10, 817−827.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., Bonas, U., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326, 1509−1512.

Bonas, U., Stall, R.E., Staskawicz, B., 1989. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. vesicatoria. Mol. Gen. Genet. 218, 127−136.

Brockdorff, N., 2013. Noncoding RNA and polycomb recruitment. RNA 19, 429−442.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., van der Oost, J., 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960−964.

Bulyk, M.L., Johnson, P.L., Church, G.M., 2002. Nucleotides of transcription factor binding sites exert interdependent effects on the binding affinities of transcription factors. Nucleic Acids Res. 30, 1255−1261.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., Voytas, D.F., 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 39, e82.

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., Jaenisch, R., 2013.

Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res. 23, 1163−1171.
Choo, Y., Klug, A., 1994a. Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. Proc. Natl. Acad. Sci. USA 91, 11168−11172.
Choo, Y., Klug, A., 1994b. Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. Proc. Natl. Acad. Sci. USA 91, 11163−11167.
Choulika, A., Perrin, A., Dujon, B., Nicolas, J.F., 1995. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. Mol. Cell. Biol. 15, 1968−1973.
Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819−823.
Dahlem, T.J., Hoshijima, K., Jurynec, M.J., Gunther, D., Starker, C.G., Locke, A.S., Weis, A.M., Voytas, D.F., Grunwald, D.J., 2012. Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. PLoS Genet. 8, e1002861.
Davidovich, C., Zheng, L., Goodrich, K.J., Cech, T.R., 2013. Promiscuous RNA binding by polycomb repressive complex 2. Nat. Struct. Mol. Biol. 20, 1250−1257.
Del Rio, S., Menezes, S.R., Setzer, D.R., 1993. The function of individual zinc fingers in sequence-specific DNA recognition by transcription factor IIIA. J. Mol. Biol. 233, 567−579.
DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M., 2013. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 41, 4336−4343.
Dickinson, D.J., Ward, J.D., Reiner, D.J., Goldstein, B., 2013. Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. Nat. Methods 10, 1028−1034.
Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., Church, G.M., 2013. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat. Methods 10, 1116−1121.
Farzadfard, F., Perli, S.D., Lu, T.K., 2013. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. ACS Synth. Biol. 2, 604−613.
Flockhart, I., Booker, M., Kiger, A., Boutros, M., Armknecht, S., Ramadan, N., Richardson, K., Xu, A., Perrimon, N., Mathey-Prevot, B., 2006. FlyRNAi: the *Drosophila* RNAi screening center database. Nucleic Acids Res. 34, D489−D494.
Flockhart, I.T., Booker, M., Hu, Y., McElvany, B., Gilly, Q., Mathey-Prevot, B., Perrimon, N., Mohr, S.E., 2012. FlyRNAi.org − the database of the *Drosophila* RNAi screening center: 2012 update. Nucleic Acids Res. 40, D715−D719.
Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., Sander, J.D., 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. 31, 822−826.
Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468, 67−71.
Gasiunas, G., Barrangou, R., Horvath, P., Siksnys, V., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA 109, E2579−E2586.
Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442−451.
Gloor, G.B., Nassif, N.A., Johnson-Schlitz, D.M., Preston, C.R., Engels, W.R., 1991. Targeted gene replacement in *Drosophila via* P element-induced gap repair. Science 253, 1110−1117.
Golic, K.G., Lindquist, S., 1989. The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome. Cell 59, 499−509.
Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., O'Connor-Giles, K.M., 2013. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. Genetics 194, 1029−1035.
Grissa, I., Vergnaud, G., Pourcel, C., 2007a. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics 8, 172.
Grissa, I., Vergnaud, G., Pourcel, C., 2007b. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 35, W52−W57.
Grissa, I., Vergnaud, G., Pourcel, C., 2008. CRISPRcompar: a website to compare clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 36, W145−W148.
Groth, A.C., Fish, M., Nusse, R., Calos, M.P., 2004. Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. Genetics 166, 1775−1782.
Gupta, A., Hall, V.L., Kok, F.O., Shin, M., McNulty, J.C., Lawson, N.D., Wolfe, S.A., 2013. Targeted chromosomal deletions and inversions in zebrafish. Genome Res. 23, 1008−1017.
Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., Cradick, T.J., Marraffini, L.A., Bao, G., Zhang, F., 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol 31, 827−832.
Huang, J., Zhou, W., Watson, A.M., Jan, Y.N., Hong, Y., 2008. Efficient ends-out gene targeting in *Drosophila*. Genetics 180, 703−707.
Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., Joung, J.K., 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. 31, 227−229.
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. J. Bacteriol. 169, 5429−5433.
Jansen, R., Embden, J.D., Gaastra, W., Schouls, L.M., 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. 43, 1565−1575.
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816−821.
Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., Doudna, J., 2013. RNA-programmed genome editing in human cells. Elife 2, e00471.
Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., Regev, A., Lander, E.S., Rinn, J.L., 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc. Natl. Acad. Sci. USA 106, 11667−11672.
Kondo, S., Ueda, R., 2013. Highly improved gene targeting by germline-specific Cas9 expression in *Drosophila*. Genetics 195, 715−721.
Lange, S.J., Alkhnbashi, O.S., Rose, D., Will, S., Backofen, R., 2013. CRISPRmap: an automated classification of repeat conservation in prokaryotic adaptive immune systems. Nucleic Acids Res. 41, 8034−8044.
Lee, H.Y., Haurwitz, R.E., Apffel, A., Zhou, K., Smart, B., Wenger, C.D., Laderman, S., Bruhn, L., Doudna, J.A., 2013. RNA-protein analysis using a conditional CRISPR nuclease. Proc. Natl. Acad. Sci. USA 110, 5416−5421.
Li, T., Huang, S., Zhao, X., Wright, D.A., Carpenter, S., Spalding, M.H., Weeks, D.P., Yang, B., 2011. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res. 39, 6315−6325.
Lin, M.F., Carlson, J.W., Crosby, M.A., Matthews, B.B., Yu, C., Park, S., Wan, K.H., Schroeder, A.J., Gramates, L.S., St Pierre, S.E., Roark, M., Wiley Jr., K.L., Kulathinal, R.J., Zhang, P., Myrick, K.V., Antone, J.V., Celniker, S.E., Gelbart, W.M., Kellis, M., 2007. Revisiting the protein-coding gene catalog of *Drosophila* melanogaster using 12 fly genomes. Genome Res. 17, 1823−1836.
Liu, J., Li, C., Yu, Z., Huang, P., Wu, H., Wei, C., Zhu, N., Shen, Y., Chen, Y., Zhang, B., Deng, W.M., Jiao, R., 2012. Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. J. Genet. Genomics 39, 209−215.
Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., Church, G.M., 2013a. CAS9 transcriptional activators for target

specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. 31, 833−838.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013b. RNA-guided human genome engineering *via* Cas9. Science 339, 823−826.

McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P., Hirschhorn, J.N., 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat. Rev. Genet. 9, 356−369.

Meader, S., Ponting, C.P., Lunter, G., 2010. Massive turnover of functional sequence in human and other mammalian genomes. Genome Res. 20, 1335−1343.

Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., Almendros, C., 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology 155, 733−740.

Moscou, M.J., Bogdanove, A.J., 2009. A simple cipher governs DNA recognition by TAL effectors. Science 326, 1501.

Park, P.J., 2009. ChIP-seq: advantages and challenges of a maturing technology. Nat. Rev. Genet. 10, 669−680.

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., Guilak, F., Crawford, G.E., Reddy, T.E., Gersbach, C.A., 2013. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat. Methods 10, 973−976.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173−1183.

Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., Zhang, F., 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154, 1380−1389.

Ren, X., Sun, J., Housden, B.E., Hu, Y., Roesel, C., Lin, S., Liu, L.P., Yang, Z., Mao, D., Sun, L., Wu, Q., Ji, J.Y., Xi, J., Mohr, S.E., Xu, J., Perrimon, N., Ni, J.Q., 2013. Optimized gene editing technology for *Drosophila melanogaster* using germ line-specific Cas9. Proc. Natl. Acad. Sci. USA 110, 19012−19017.

Reyon, D., Tsai, S.Q., Khayter, C., Foden, J.A., Sander, J.D., Joung, J.K., 2012. FLASH assembly of TALENs for high-throughput genome editing. Nat. Biotechnol. 30, 460−465.

Rong, Y.S., Golic, K.G., 2000. Gene targeting by homologous recombination in *Drosophila*. Science 288, 2013−2018.

Rousseau, C., Gonnet, M., Le Romancer, M., Nicolas, J., 2009. CRISPI: a CRISPR interactive database. Bioinformatics 25, 3317−3318.

Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., Lin, M.F., Washietl, S., Arshinoff, B.I., Ay, F., Meyer, P.E., Robine, N., Washington, N.L., Di Stefano, L., Berezikov, E., Brown, C.D., Candeias, R., Carlson, J.W., Carr, A., Jungreis, I., Marbach, D., Sealfon, R., Tolstorukov, M.Y., Will, S., Alekseyenko, A.A., Artieri, C., Booth, B.W., Brooks, A.N., Dai, Q., Davis, C.A., Duff, M.O., Feng, X., Gorchakov, A.A., Gu, T., Henikoff, J.G., Kapranov, P., Li, R., MacAlpine, H.K., Malone, J., Minoda, A., Nordman, J., Okamura, K., Perry, M., Powell, S.K., Riddle, N.C., Sakai, A., Samsonova, A., Sandler, J.E., Schwartz, Y.B., Sher, N., Spokony, R., Sturgill, D., van Baren, M., Wan, K.H., Yang, L., Yu, C., Feingold, E., Good, P., Guyer, M., Lowdon, R., Ahmad, K., Andrews, J., Berger, B., Brenner, S.E., Brent, M.R., Cherbas, L., Elgin, S.C., Gingeras, T.R., Grossman, R., Hoskins, R.A., Kaufman, T.C., Kent, W., Kuroda, M.I., Orr-Weaver, T., Perrimon, N., Pirrotta, V., Posakony, J.W., Ren, B., Russell, S., Cherbas, P., Graveley, B.R., Lewis, S., Micklem, G., Oliver, B., Park, P.J., Celniker, S.E., Henikoff, S., Karpen, G.H., Lai, E.C., MacAlpine, D.M., Stein, L.D., White, K.P., Kellis, M., 2010. Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. Science 330, 1787−1797.

Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K., Dobbs, D., 2010. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. Nucleic Acids Res. 38, W462−W468.

Sander, J.D., Zaback, P., Joung, J.K., Voytas, D.F., Dobbs, D., 2007. Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. Nucleic Acids Res. 35, W599−W605.

Sanjana, N.E., Cong, L., Zhou, Y., Cunniff, M.M., Feng, G., Zhang, F., 2012. A transcription activator-like effector toolbox for genome engineering. Nat. Protoc. 7, 171−192.

Sebo, Z.L., Lee, H.B., Peng, Y., Guo, Y., 2013. A simplified and efficient germline-specific CRISPR/Cas9 system for *Drosophila* genomic engineering. Fly (Austin) 8.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., Zhang, F., 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343, 84−87.

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J.L., Gao, C., 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. 31, 686−688.

Skennerton, C.T., Imelfort, M., Tyson, G.W., 2013. Crass: identification and reconstruction of CRISPR from unassembled metagenomic data. Nucleic Acids Res. 41, e105.

Smih, F., Rouet, P., Romanienko, P.J., Jasin, M., 1995. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. Nucleic Acids Res. 23, 5012−5019.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., Jaenisch, R., 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153, 910−918.

Wang, T., Wei, J.J., Sabatini, D.M., Lander, E.S., 2014. Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80−84.

Wei, C., Liu, J., Yu, Z., Zhang, B., Gao, G., Jiao, R., 2013. TALEN or Cas9 - rapid, efficient and specific choices for genome modifications. J. Genet. Genomics 40, 281−289.

Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X., Zhu, Z., Lin, S., Zhang, B., 2013. Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res. 41, e141.

Yu, Z., Ren, M., Wang, Z., Zhang, B., Rong, Y.S., Jiao, R., Gao, G., 2013. Highly efficient genome modifications mediated by CRISPR/Cas9 in *Drosophila*. Genetics 195, 289−291.

Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., Lee, J.T., 2010. Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol. Cell 40, 939−953.