Review

Engineering the *Drosophila* Genome for Developmental Biology

Dagmara Korona, Stefan A. Koestler 🆔 and Steven Russell *

Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK;
dk500@cam.ac.uk (D.K.); sak76@cam.ac.uk (S.A.K.)
* Correspondence: s.russell@gen.cam.ac.uk; Tel.: +44-1223-766-929

Abstract: The recent development of transposon and CRISPR-Cas9-based tools for manipulating the fly genome in vivo promises tremendous progress in our ability to study developmental processes. Tools for introducing tags into genes at their endogenous genomic loci facilitate imaging or biochemistry approaches at the cellular or subcellular levels. Similarly, the ability to make specific alterations to the genome sequence allows much more precise genetic control to address questions of gene function.

Keywords: *Drosophila*; genome engineering; CRISPR-Cas9; protein tagging

1. Introduction

It is over a century since Thomas Hunt Morgan and his students established *Drosophila melanogaster* as a powerful system for exploring the basis of heredity [1,2]. By the latter period of the 20th century, due in part to the pioneering work of Ed Lewis, along with the large scale genetics screens by Christiane Nüsslein-Volhard and Eric Wieschaus, and coupled with the tremendous progress in molecular biology, the fly became well established as a model for developmental biology [2]. With an increasingly sophisticated range of genetic tools, very efficient transgenesis and the well-described anatomy available, considerable progress was made in characterising conserved developmental pathways [3]. The advent of genome sequencing and the post-genomic era inevitably led to new approaches in developmental biology, in particular, functional analysis of genes identified in high-throughput genomics screens or fly orthologues of mammalian genes identified in such studies [4]. Despite its sophistication, the major drawback of the fly as an experimental model was the lack of an easy system for targeting genomic changes in the organism via the type of homologous recombination used in yeast or vertebrate cells. While the development of gene targeting approaches [5,6] by the Golic laboratory addressed this deficiency in part, the systems are cumbersome and success tends to be highly locus specific. More recently, a range of transposon-based methods has allowed much greater flexibility in precisely manipulating genes in situ [7,8], however, the very recent advent of Clustered, Regularly Interspaced, Short Palindromic Repeat (CRISPR)-based genome engineering now offers the prospect of making virtually any desired change to the genome sequence much easier than previously achievable. Together, these methods are opening new avenues in developmental biology research, facilitating precise questions about gene/protein function, allowing the targeted expression or knock out of any gene and enabling high-resolution imaging of proteins, often in live tissues. Here we overview the most recent methods for manipulating the fly genome with a focus on approaches tackling gene manipulation and in vivo localisation for developmental studies.

2. Classical Methods

One of the major landmarks in the progression of the fly as an experimental tool for the study of development was the establishment of a robust method for generating transgenic animals via the *P* element transposon [9]. Since its introduction, a range of sophisticated *P* element-based tools have been constructed; including the Gal4-*UAS* system for precise spatial and temporal control of gene expression [10]; enhancer trapping for the genome-wide identification of developmentally regulated genes [11,12] and methods for systematic gene misexpression with engineered *P* transposable elements (EP) [13]. The utility and use of these by now classic tools have been well described in many reviews, and we here focus on some of the more modern applications of transposons. While the *P* element has been the workhorse of *Drosophila* transgenesis for over 30 years, it does exhibit an insertion bias [14], particularly towards gene promoters, making it less useful for targeting other genic regions such as introns. The establishment of other transposon systems, particularly *Minos* and *piggyBac* [15,16], along with the site-specific integrase PhiC31 [17], have opened more of the fly genome to engineering approaches. Considerable detail on newer engineering methods have recently been well reviewed by Venken and colleagues [18]; here we briefly touch on the most important systems, with a focus on those facilitating imaging studies or introducing precise changes to the genome sequence.

The introduction of a variety of inducible site-specific recombination systems, principally Flp-*FRT* [19] and Cre-*Lox* [20], have enhanced our ability to manipulate gene expression in space and time or to precisely engineer the genome. For example, introducing transgenic constructs containing *FRT* sites into the genome has considerable utility for functional and imaging analyses, facilitating spatial or temporal control of gene activation [21]. Generating mitotic or germ-line clones, techniques that are necessary to overcome pleiotropic or maternal effects of mutations and study gene functions during later stages of development, are simplified by the use of chromosomes with *FRT* sites close to the centromere. Similarly, the use of Flp-mediated recombination sees widespread use in methods aimed at marking specific populations of cells for lineage tracing or functional studies, particularly in the nervous system, and these have recently been well reviewed by others [22–27].

Projects by Exelixis and the DrosDel consortium have generated several thousand insertions with transposons carrying *FRT* sites that allow precise chromosome engineering [28,29]. For example, using the approach developed by Golic and Golic [30] (Figure 1a–c), DrosDel generated over 3000 insertions of re-arrangement screen 5 and 3 (RS5 and RS3) *P* elements, using these to build deletion coverage for almost 80% of the genome [29]. More importantly, combining the Exelixis and the DrosDel collections allows the construction of over 500,000 deletions ranging in size from 1 bp to 1 Mb that are precisely mapped at base pair resolution. These tools can be used to construct precise deletions in a homogeneous genetic background, for example, facilitating the analysis of complex regulatory regions [31], or genome-wide analysis of gene dosage effects to provide insights into gene regulatory network robustness [32]. Finally, *FRT* elements may be combined to generate other types of chromosomal aberrations, including translocations and inversions [30], and we have used such an approach to examine the consequences of disrupting developmental gene regulatory neighbourhoods [33]. The ability to precisely manipulate genomes at the chromosome level remains a powerful tool in the research armoury and the Bloomington Deficiency kit, rebuilt using the Flp-*FRT* approach, remains a widely used resource for the type of gene mapping and pathway expansion studies that have been important in building developmental networks [34].

The phiC31 system is a particularly useful tool for generating transgenic flies since a range of different transgenes may be inserted at an identical site in the genome [17,35], ameliorating concerns about position effects or other insertion site artefacts when comparing the effects of different constructs. PhiC31 integrase is directional and highly specific, mediating recombination between *attP* and *attB* sites to generate hybrid *attL* and *attR* sites that are not substrates for the enzyme (Figure 1d). A set of lines containing landing sites across the genome are available from stock centres and at least some of these have been functionally characterised in terms of expression properties [36]. Recently, a mutated integrase has been developed that allows reversal of the integration reaction [37], which

is useful for building arrays of different transgenes at single sites in the genome. A second very useful approach available with phiC31 is recombinase-mediated cassette exchange (RMCE) [38]: in this case a genomic landing site containing a marker gene flanked by *attP* sites may be replaced by any other DNA sequence via a plasmid containing a gene or sequence of interest flanked by *attB* sites (Figure 1e). The RMCE system is incredibly powerful and since its introduction has been deployed as a component in a wide variety of genome engineering strategies (i.e., see Minos-mediated integration cassette (MiMIC) below).
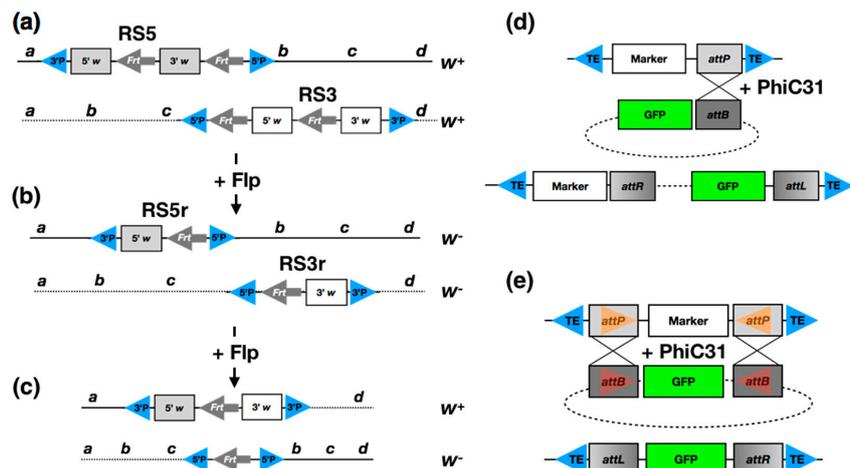

Figure 1. Genome engineering. (**a–c**) **The Golic method for generating precise chromosomal deletions.** (**a**) Two *P* elements, designated RS3 and RS5, are inserted at different locations (designated by a–d) on two homologous chromosomes and kept in separate fly lines. The elements contain a functional mini-*white* gene composed of multiple exons that for simplicity are drawn as grey or white boxes, representing 5′ and 3′ portions of the gene. There are two *Frt* sites (grey arrows) in each element, one of which is located within a mini-*white* exon. The elements differ in the position of the second *Frt* site and the orientation of the construct with respect to the *P* element ends (blue triangles). (**b**) Internal Flp-driven recombination between the *Frt* sites produces remnant forms of the *white* genes such that RS5r contains the 5′ end and RS3r the 3′ end with the intronic *FRT* site remaining. Each of these remnant elements are generated in separate fly lines that are phenotypically white eyed. (**c**) RS5r and RS3r elements are brought together in trans in a fly along with a source of Flp recombinase. FLP-mediated recombination between the elements produces a reconstituted functional *white* gene and the intervening genomic DNA is deleted. The reciprocal event creates a tandem duplication of the deleted segment, separated by an *FRT* site, but no white gene. (**d**) **The PhiC31 system.** A transposon (blue triangles mark the transposon ends) carrying a marker gene for genetic tracking and an *attP* site is inserted into the genome. Providing an *attB* containing plasmid with a gene or sequence of interest, in this case GFP, and a source of PhC31 integrase results in high efficiency integration of the plasmid into the genomic location. (**e**) **Recombinase-mediated cassette exchange (RMCE).** A transposon (blue triangles mark the transposon ends) carrying a marker gene flanked by *attP* site is inserted into the genome. Providing a plasmid with a gene or sequence of interest, in this case GFP, flanked by *attB* sites and a source of PhC31 recombinase results in high efficiency replacement of the genomic marker with the sequence of interest. With RMCE the inverted orientation of the *attP* and *attB* sites is critical for producing the desired exchange.

A key requirement for understanding gene function during development is the ability to localise gene products in space and time at the organismal, cellular and subcellular levels. For many years immunohistochemistry or immunofluorescence using specific antibodies was the primary method for collecting these data. Where antibodies are not available or there is a requirement for live imaging, transgenes containing tagged versions of a protein of interest have been useful. It is most common to express genes encoding tagged proteins under the control of the Gal4-*UAS* system, however, there can

be concerns that over or misexpression may affect function. While it is sometimes possible to generate tagged transgenes that rescue null mutants, endogenous regulatory sequences are often too large or poorly characterised to make such an approach generally applicable. To try and address this P[acman] bacterial artificial chromosome (BAC) libraries for *D. melanogaster* [39] and Fosmid libraries for both *D. melanogaster* and *D. pseudoobscura* have been generated, which allow recombineering approaches in *E. coli* to generate tagged transgenes that are then integrated into the fly genome [40,41]. It would however clearly be preferable to tag endogenous loci in situ. Two transposon based approaches have been developed to facilitate in vivo protein tagging: the MiMIC method described below and various protein trap strategies. In the latter, a transposon containing an artificial exon, usually a fluorescent protein flanked by splice acceptor and donor sequences, is mobilised in the germline and progeny inheriting fluorescent reporter expression are recovered. A comprehensive screen requires a set of three vectors, each with the reporter in a different open reading frame. When the transposon inserts into an intron of a protein coding gene in the correct orientation, splicing will incorporate the artificial exon into the primary transcript and hence the fluorescent reporter into the protein [42] (Figure 2a–c). Due to the bias observed with *P* elements, particularly insertion hotspots in promoter and 5′ UTR regions [43], the use of *piggyback*-based transposons was found to generate a much higher frequency of bona fide protein traps. Together, three large-scale studies screened approximately 150 million embryos, recovering a little over 600 verified protein traps between them. Reassuringly, the majority of protein traps do not appear to substantially affect protein function [44–46]. While this may seem a relatively poor yield for such an effort, the lines generated have been incredibly useful for imaging studies down to the subcellular level [47] and have also facilitated mass spectrometry-based analysis of in vivo protein complexes that are purified via the tag [48]. Providing the ability to characterise protein localisation and interactions, protein trap lines have been used to study diverse developmental processes, including live imaging of anterior-posterior (A-P) polarity formation during oogenesis, the cell biology of epithelial morphogenesis and signalling pathway interactions [49–51]. Trapped lines have also facilitated the characterisation of genome wide transcription factor binding via chromatin immunoprecipitation (ChIP)-based studies [52].

One of the most useful and flexible transposon-based strategies available to the fly biologist is the MiMIC system [53] (Minos-mediated integration cassette): a collection of over 7000 *Minos*-based insertions that provide a platform for a variety of targeted engineering approaches. The core MiMIC transposon contains inverted phiC31 *attP* target sites within the *Minos* ends, a splice acceptor site followed by stop codons in all three reading frames. The element also contains a transcriptional/translational cassette encoding EGFP followed by a polyadenylation signal and a phenotypic marker (*yellow*). Insertion of a MiMIC element in the correct orientation into an intron of a coding gene will generate a truncated protein due to the splice acceptor and stop codons, thus acting as a gene trap. The true utility of MiMIC is the ability to use the phiC31 *attP* sites to replace the contents of the transposon with any piece of DNA via the RMCE system (Figure 2d). For example, regulatory sequences may be added to genes, functional reporters such as Gal4 or Flp may be inserted within 5′ UTRs or, in cases when inserts are within coding introns, a wide range of protein tags may be introduced. Around 30 different cassettes have been generated by the MiMIC team (Figure 2e), but any desired sequence can be easily cloned and introduced into the genome using the system. More recently, the team developed a genetic approach to circumvent the requirement for embryo injection [54], speeding up the generation of cassette replacement. Approximately 2000 genes currently have MiMIC insertions within introns, but the possibility of using CRISPR (see below) to place MiMIC insertions into the genome opens up the possibility that every intron containing gene in the genome could be amendable to protein trapping. The utility of the MiMIC system is attested to by over 100 research papers citing its use in a little over six years.
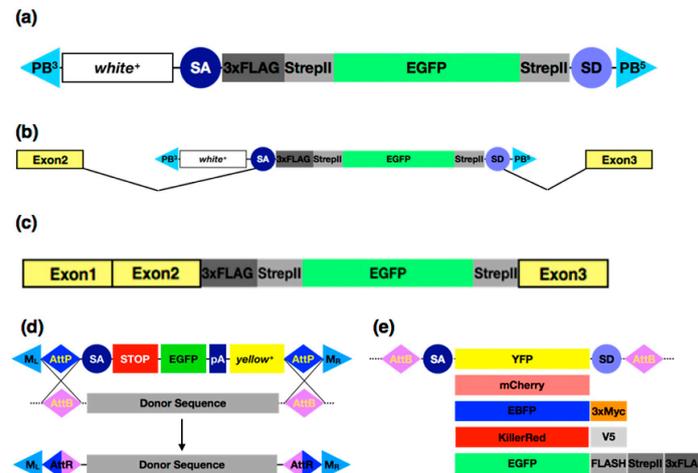
Figure 2. Protein Trapping. (**a**–**c**) *piggyBac* **transposon-based protein trapping**. (**a**) The *pigP* protein trap element used in [47]: the transposon ends (blue triangles) flank a genetic marker gene (*white*) and an artificial exon, which in this case contains the coding sequence for enhanced green fluorescent protein (EGFP) along with StrepII and 3XFLAG tags between splice donor and splice acceptor sites (blue circles). (**b**) Insertion of the transposon into an intron of a protein coding gene (represented by lines separating the yellow boxes) allows the possibility of splicing the artificial exon into the gene transcript. (**c**) If the transcript carrying the artificial exon is translated a tagged protein is generated. (**d**,**e**) **The** Minos-mediated integration cassette **(MiMIC) System.** (**d**) A *Minos*-based transposon with transposable element (TE) ends indicated by blue triangles, contains two *attP* sites in inverted orientation (blue diamonds) flanking a gene trap cassette with a splice acceptor (blue circle), stop codons in all three reading frames (red box), a fluorescent marker (EGFP) followed by a polyadenylation signal (blue box) and a genetic marker (*yellow*). The sequences internal to the *attP* sites may be replaced via a RMCE reaction by providing a donor sequence flanked by *attB* sites and a source of phiC31 integrase. (**e**) A variety of different fluorescent reporters have been developed that can be used to introduce tags into genes with MiMIC insertions in coding introns via RMCE.

Whether protein trap or engineered MiMIC strategies are used to generate tagged proteins, there are some useful approaches to knock down their function [55] in a tissue or temporally specific manner. Transcripts containing inserted tags may be targeted by RNA-mediated interference (RNAi) constructs directed against the tag, and lines containing *UAS*-driven RNAi recognising green-fluorescent protein/enhanced green fluorescent protein (GFP/EGFP) are available from stock centres. A second approach is via the use of an anti-GFP nanobody, which targets GFP-tagged proteins for degradation via the ubiquitin pathway [56,57] and *UAS*-driven lines are available. Both these methods have been shown to be effective with in vivo tagged proteins and offer a level of functional control over gene products that can, in principle, be monitored in real time via imaging of the fluorescent tag. However, despite the relative ease and simplicity of these methods they have obvious limitations; the elimination of tagged proteins is relatively slow and they are not readily reversible. For the rapid elimination of gene function, an auxin-inducible degradation system (AID) has recently been shown to work in *Drosophila* [58]. This system offers the ability to regulate the inducible and rapid degradation of specific target proteins by the recruitment of proteins with an auxin-inducible degron to the E3 ubiquitin ligase complex. A disadvantage of this system is the need to introduce a specific tag containing an auxin-inducible degron domain into the gene of interest, though this could be achieved for genes harbouring an appropriate MiMIC insertion.

3. CRISPR/Cas9

3.1. Overview

As the above discussion emphasises, the range of transposon based tools available to the *Drosophila* biologist is impressive, however, even here the landscape for genome engineering has changed dramatically with the introduction of the CRISPR-Cas9 system. Clustered, Regularly Interspaced, Short Palindromic Repeat (CRISPR)-based engineering has become the most popular and commonly used tool for genome manipulation over the last five years. Since its discovery, or rather understanding its potential, numerous studies have confirmed the wide applicability and adaptability of the system [59–61] and a number of different methods have been developed with different variants of the RNA-guided Cas9 endonuclease adopted for genome editing [62].

CRISPR-Cas systems are natural RNA-guided adaptive immune systems that provide protection mechanisms against viruses or other exogenous DNA entering bacterial and archaeal cells. The most studied and widely applied system for genome editing, Class 2 CRISPR-Cas from *Streptococcus pyogenes*, uses a guide RNA (gRNA) composed of a specific CRISPR RNA (crRNA) combined with a tracrRNA sequence to guide the Cas9 nuclease to a genomic target site where it introduces a double-strand break (DSB) that is repaired by endogenous repair pathways. The system requires only 20 nucleotides of homology to the target site in the gRNA, but the target must be adjacent to a three nucleotide (NGG) protospacer adjacent motif (PAM) in the genome (Figure 3a) [63,64]. Repair may be via the non-homologous end joining (NHEJ) pathway [65], leading to insertions and/or deletions (indels) at the target site. Alternatively, the activity of homology-directed repair (HDR) pathways can facilitate precise genome editing by introducing alternative sequences via a supplied donor DNA (Figure 3a) [66,67]. While the initial efficiency of specific targeting, especially when trying to introduce mutations or tags, was initially low and variable [68,69], various improved systems have been developed. The method originally relied on delivery of the nuclease and RNA components via microinjection of RNA or DNA, the development of transgenic lines expressing Cas9 under the control of germline promoters such as *nanos* (*nos*) or *vasa* as recipients for injection of plasmids expressing gRNAs from U6 promoters has considerably simplified the process and can generate desired events in the genome with higher efficiency than with the initial injection approach [70].

One of the major concerns in any genome engineering work, whether using CRISPR/Cas9 or any other system, is the possibility of off-target effects resulting from gRNA binding to imperfectly matched targets at other genomic locations [71]. While the extent of off-target effects has not been fully characterised, several studies that suggest that for relatively small genomes, such as that of *Drosophila*, off-targets may be less of an issue compared to organisms with larger genomes [72,73]. The problem may be, at least in part, overcome by using a mutant version of Cas9 developed by Cong and colleagues [74], Cas9D10A (Cas9n), that makes single strand nicks at the target locus [74–76]. When used in conjunction with two gRNAs flanking the desired target site and a donor repair template, repair is generally via the high-fidelity HDR pathway [69,77]. The nickase activity of Cas9n does not activate the NHEJ pathway and while Cas9n is certainly more specific than wild-type Cas9, DSBs are still detectable at target sites when only one gRNA is used and it is possible that Cas9n may cause indels at off-target sites due to either of the gRNAs. Encouragingly, the paired Cas9n system has been used to generate large deletions without observing unwanted translocations, and so forth (Figure 3b) [76,78], and a recent comparative study indicates that Cas9n can induce more HDR than NHEJ, at least in human HEK293 cells [79].

Concerns regarding the use of Cas9n have been raised due to the possibility of DNA rearrangements via single nicks [62]. These unwanted off-targets can result from the recognition of multiple genomic sites by gRNAs that that can subsequently be cleaved. In contrast to other nuclease systems for genome engineering, such as zinc finger nucleases or transcription activator-like effector nucleases (TALENs), that rely on dimerization for target cleavage, Cas9 acts as a monomer. In order to increase specificity, new Cas9 versions that only cleave when dimerised have been developed:

for example, a nuclease dead Cas9 (dCas9) fused to the non-specific endonuclease FokI has the gRNA directed specificity of the CRISPR system but relies on obligate FokI dimerization for cleavage (Figure 3c) [63,64,80,81]. A clear limitation of the paired dCas9-FolkI approach is the requirement for two appropriately spaced gRNA sites 15–25 bases apart [81]. This constrains the widest applicability of the method since these conditions cannot always be achieved. Alternatively, there have been efforts to develop engineered Cas9 enzymes (i.e., SpCas9-HF1) with increased fidelity to ensure more specific targeting [82], however, there are reports of lower efficiencies with such variants [83]. Irrespective of the engineering method used, effects of off-target events may be largely ameliorated by cleaning up stocks via standard back-crossing methods.
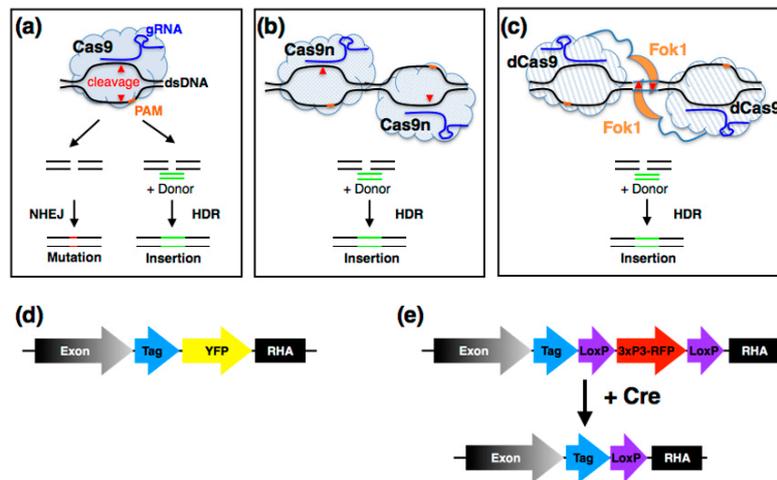

Figure 3. CRISPR Genome Engineering. (**a**) The wild-type Cas9 complex (light blue cloud) contains the Cas9 endonuclease and a guide RNA (gRNA) (blue) complementary to the target site adjacent to a PAM sequence (orange). The complex opens the DNA duplex and introduces a double strand break (red triangles). Repair by the non-homologous end joining (NHEJ) pathway may result in indel mutations whereas homology directed repair (HDR) in the presences of a donor template (green) generates insertions. (**b**) Using mutant Cas9n enzymes that make single strand cuts with two gRNAs (blue) direct the Cas9n complexes to make cuts (red triangles) separated by some distance. The gap may be repaired in the presence of a donor (green) to generate an insertion. (**c**) Dead Cas9 (dCas9) enzymes (which are unable to cleave DNA) are fused with FokI nuclease monomers (orange). When two gRNAs (blue) some distance apart are used, dCas9-FolkI monomers are brought into proximity allowing the FokI to dimerise and cleave in between. The resulting gap may be repaired in the presence of a donor (green) to generate an insertion. (**d**) Diagram of donor cassette for direct addition of a protein tag to the C-terminal of a coding exon. The donor DNA contains part of the exon sequence with a biochemistry tag and yellow fluorescent protein (YFP) in frame, followed by a right hand end homology arm (RHA) that can mediate the type of insertion event shown in (**a**). (**e**) Diagram of a donor sequence used to introduce biochemistry tags to the C-terminal of a coding exon along with a removable marker. The donor DNA contains part of the exon sequence with a biochemistry tag in frame. This is followed by an eye expressed red fluorescent protein (RFP) cassette flanked by *LoxP* sites, which are downstream of the splice donor site and thus within an intron, then a right hand homology arm. Once a verified insertion has been recovered, tracked by the acquisition of RFP, the RFP is removed by exposing to Cre recombinase and recovering flies who have lost RFP expression. The resulting lines have the tagged exon and a single intronic *LoxP* site.

The obvious restriction when using Cas9 for genome engineering is the requirement for a PAM sequence adjacent to the desired target site, which limits the number of Cas9 engineerable sites in a genome. To overcome this limitation several mutant versions of Cas9 with altered specificities have been engineered [84] and Cas9 homologs from other organisms that show different PAM

specificities have been isolated. For example, a recently identified class 2 endonuclease, Cpf1 from *Francisella novicida*, has received some attention [85]. The Cpf1 PAM is TTN and is located 5′ to the target site (in contrast to the Cas9 PAM which is 3′ to the target), providing a new set of potential targets not accessible to Cas9 systems. Finally, other Class 2 CRISPR effectors, such as Cas13a (C2C2), have been shown to target RNA, however, few studies have explored the potential of this activity [62,86]. It should be remembered that the potential of CRISPR based systems has only been apparent for five years and the progress in developing variants systems with different properties has been tremendous and will surely continue to provide new enzymes and systems with improved specificity and fidelity.

3.2. Mutagenesis with CRISPR/Cas9 Systems

The most straight forward way to modify the genome using CRISPR/Cas9 based genome engineering is the introduction of short insertions/deletions (indels) by inducing NHEJ, which frequently leads to frameshift mutations and thus null alleles or truncations of the targeted gene product. All that is required is the Cas9 protein and a single gRNA targeting the gene of interest after the transcription start site [87]. Large deletions of 30 Mb have been generated in human cell lines and even chromosomal rearrangements such as translocations can be generated by providing two gRNAs, targeting either side of the region to be deleted or sites on two different chromosomes [87–89]. Insertions of up to 5.7 kb have been achieved by providing linear DNA during CRISPR/Cas9 induced NHEJ in Zebrafish [90], and up to 34 kb in human cell lines [91] using *piggyBac* or adenoviral vectors that were cleaved by Cas9, targeted via a vector specific gRNA. A drawback is that the position of the knock-in is not precisely defined, and thus a molecular screen for in-frame insertions must be performed. It remains to be determined to what extent NHEJ-driven rearrangement can be exploited in *Drosophila*.

The considerable versatility of CRISPR/Cas9 based genome editing lies in the utilization of homology directed repair (HDR) for the precise incorporation of a donor sequence in the form of ssDNA-oligonucleotides, PCR (Polymerase Chain Reaction) products or plasmids, that are carried between homology arms, which are used as templates during the double strand repair process rather than the homologous chromosome [69,92]. Depending on the target site(s), the gRNA(s), the homology arms used and the region between them, different outcomes are possible. An insertion is achieved when the donor carries a sequence between the homology arms that cover the regions directly adjacent to the double strand break, thus inserting the sequence (Figure 3a). In this way, for example, additional amino acids can be added to a coding exon, point mutations introduced, or new binding sites added to a regulatory region. The optimum length for homology arms appears to be approximately 1 kb [93]. If the homology arms are separated with respect to the target genomic DNA the donor DNA between the homology arms may replace the intervening genomic segment. Since HDR efficiency decreases with the distance of homology deviations from the double strand break, replacement or deletion of large regions can be obtained by providing two gRNAs, each next to a homology arm, thus actively excising the intervening region [93–95]. Again, the sequence between the homology arms is inserted. Note that if homologous recombination does not take place the region between the homology arms may still be deleted, which obviously generates a deletion and perhaps a null allele. While a null allele can also be produced when only one gRNA is used as indicated above, this is likely to generate a frame shift mutation, which may be less useful, especially in non-coding regions. If a specific point mutation at a locus is desirable, a dCas9-cytidine deaminase chimeric enzyme may be used rather than homology directed replacement. The deaminase catalyses the conversion of C to U residues thus generating C to T or G to A substitutions. This 'base editing' is generally more efficient than HDR-mediated point mutation and at the same time minimizes indel formation [96]. This has recently been complemented by conversion of A-T to G-C base pairs by using a transfer RNA specific adenosine deaminase fused to catalytically impaired CRISPR/Cas9 [97].

Initial screening for positive events can be done using standard PCR, high resolution melt analysis (HRM), or sometimes more easily with visual markers. These can include body or eye colour markers,

or fluorescent proteins expressed, for example, in the eye via the 3XP3 promoter (Figure 3e). Such markers are usually flanked by *FRT* sites for subsequent removal to avoid interference with gene function or other unphysiological responses [98]. Whichever method is used, any mutations should be verified by sequencing. Experiments ranging from the generation of specific single base pair mutations to replacement of whole genes have been described [99–101]. The latter is especially relevant for evo-devo studies, for example, to define the degree of functional redundancy between paralogous or orthologous genes by swapping the coding sequences of one gene into the genomic regulatory context of another. An example of such an approach is a swap of *Sox3* with *Sox2* coding sequences in mouse, facilitated by the simple genomic architecture of this class of genes since they only contain a single exon [102]. We have recently used a similar approach with the fly orthologues of these genes (S. A. Koestler and S. Russell, unpublished). Such experiments become more complex for larger genes, which may consist of several exons separated by introns, especially if the donor or host genes have intronic regulatory sequences. In such cases addition of a cDNA to functionally replace a null allele at a donor locus may be considered, although it is probably prudent to first test whether the coding sequence of the host gene is sufficient to effect rescue. For example, using this approach, gene and domain swaps in *D. melanogaster* as well as between *D. melanogaster* and *Tribolium castaneum* have led to the dissection of the regulatory function and evolution of members of the *robo* gene family [103,104]. However, fully understanding redundancy requires the removal of all redundant members of a gene family [105]. While such efforts previously required crossing or recombination of single mutants, or removal of large chromosomal regions [106,107], CRISPR/Cas9 allows a much more precise and efficient approach. Encouragingly, as the requirement for ever more high-throughput analysis is increasing, up to four genes have already been targeted in parallel by multiplexing [108,109].

The major challenges currently encountered when applying CRISPR/Cas systems for generating mutations arise from the relatively low efficiency of the desired genome modifications and the lack of an obvious phenotype in heterozygous F1 progeny, hence the need for time-consuming molecular screening of positive variants mentioned above. Due to the laborious nature of such screens and the need to examine many progeny, several groups developed methods facilitating more efficient engineering and hence more rapid screening. Bullock and colleagues [70] developed a versatile toolbox for *Drosophila* genome engineering consisting of a set of evaluated transgenic Cas9 lines and gRNA-expression plasmids. Their studies indicated that injecting gRNA(s) into Cas9 transgenic flies increases the overall efficiency but a fully transgenic system where the gRNA expressing construct is also inserted in the genome results in much higher efficiency. In principle, this strategy requires the generation of a transgenic fly line for every gRNA of interest that subsequently needs to be crossed to a transgenic Cas9 line. Despite being more time consuming initially, the fully transgenic system means that all transgenic flies pass on mutant alleles to their progeny, with efficiencies approaching 100% observed with the majority of tested gRNAs. Consequently, there is a significant improvement in screening efficiency since the majority of progeny have an engineered chromosome. More recently, based on the fact that that multiple CRISPR events can occur in a single cell [74,110], the co-CRISPR or co-conversion method initially developed in *C. elegans* [111,112] has been also successfully applied in *Drosophila* [113]. The method relies on simultaneously targeting a gene of interest and an *ebony* marker gene: *nos-Cas9* embryos [70] are injected with a mix of gRNAs targeting both *ebony* and the gene of interest (with or without a donor template). It is expected that in any given cell where *ebony* is mutated it is likely that the Cas9 has also been active at the gene of interest, thus progeny showing loss of *ebony* are selected for molecular analysis of the target gene [113]. In *Drosophila*, the co-CRSIPR strategy was found to be more efficient for identifying NHEJ mutagenesis events rather than HDR knock-in events. A similar approach targeting *white* as an editing marker has also been reported [114], however, it should be noted that *white* is a much more common transgenesis marker than *ebony* and thus may preclude its use in some crossing schemes. These efficiency and marker gene approaches simplify the task of screening for desired engineering events by enriching for successfully edited progeny and provide an excellent framework for further development of strategies to rapidly isolate mutations.

The generation of stable stocks carrying defined gene modifications is most often achieved with a germline expressed Cas9, usually via the *vasa* or *nanos* promoters described above [70,90,115,116]. However, in some situations modification of the germline may be either lethal or undesirable and it is therefore necessary to generate mosaic animals. These can be made by simply injecting the gRNA and donor construct (if one is used) into embryos ubiquitously expressing Cas9 or alternatively by co-injecting a source of Cas9. For the detection of successfully modified cells in mosaics the use of a marker is strongly recommended. The marker can be a fluorescent protein under control of a ubiquitous promoter carried on the injected donor construct. Although using a marker greatly facilitates the screening process it is not always possible to apply and the other approaches described in the previous section need to be employed. An alternative is to utilise the Gal4/*UAS* system to drive Cas9/gRNA expression in a specific cell type or at a particular developmental stage. Conditional mutants can be established, but the time it takes until they take affect and the turnover rates of the involved agents have to be considered [69,117]. It must be emphasized that this approach has to be used with care due to complications arising from the possible introduction of multiple mutations and cellular heterogeneity, especially when NHEJ is the desired repair event or gRNAs are not highly specific. Additional control and flexibility in the design of experiments can be obtained by combining CRISPR/Cas9 with other genome engineering techniques, for example, by insertion of *FRT* or *attP* sites [69]. Zhang et al. took this a step further by combining CRISPR/Cas9 with RMCE [118] to facilitate the analysis of several different constructs at the same locus.

A given gene product often exerts its function in a variety of different cellular contexts, that is, a transcription factor used in different cells during development, and this is reflected in the control of its expression by a combination of multiple regulatory sequences [119,120]. To probe these non-coding, regulatory regions the same CRISPR/Cas9 approaches that are being applied to coding sequences are useful. These include deletion [121], insertion or swapping of regulatory elements to study functional redundancy, for example, of distributed (primary and shadow) enhancers [122–124]. Single base pair changes are suited to determine the relevance of specific positions in transcription factor binding motifs or to alter splice sites [125,126]. The readout, however, is distinct. Regulatory sequences and their combinatorial logic can be revealed by examination of expression patterns, transcription factor binding, their DNA binding dynamics, and chromatin state or three dimensional nuclear chromatin organization [88,127,128]. Knowledge of regulatory modules also allows generation of more refined tools by driving cell type and developmental stage specific expression of CRISPR/Cas9 components [123,129]. Databases such as the super enhancer archive (http://sea.edbc.org) [130] that integrate results from various types of experiments and CRISPR/Cas9 target sites can aid in the design of experiments. Integration of all these data with fluorescent tagging of gene products at their endogenous loci (see below) will facilitate measurement of local concentrations. Together, this will significantly contribute to a more comprehensive understanding of gene regulation, which will culminate in more precise mathematical models [120].

3.3. Protein Tagging with CRISPR/Cas9 Systems

As with transposon based approaches, the use of CRISPR/Cas9 genome engineering for protein tagging is beginning to provide new tools for developmental biology. In contrast to the methods described above, the CRISPR system facilities the precise introduction of protein tag sequences at endogenous loci [70]. Thus, tagged proteins are expressed under the control of endogenous regulatory sequences, more likely to be expressed at physiological levels and, if relevant information is available, may be engineered to minimise the impact on normal protein function. Introduced protein tags may be visible fluorescent proteins, for example, YFP, GFP, mCherry, and so forth [131] (Figure 3d) or short epitope tags (FLAG, STREPII, Myc, etc.) that, due to their small size, are less likely to perturb protein function. Tagged proteins may be visualised in vivo via fluorescence microscopy or immunohistochemistry and tags may also be used in biochemical studies, for example complex purification combined with mass spectrometry [48,132].

Due to the fact that the small size of epitope tags is likely to reduce perturbation since they can be more readily incorporated into a protein scaffold, versatile protein tagging approaches with split fluorescent proteins have been developed [133]. Such self-complementing split fluorescent proteins enable live cell imaging when they are used as epitope tags and have been shown to work in *Drosophila* cells [134]. The other advantage, aside from the small tag size, is the ease with which it is adaptable to genome engineering since successful tag knock-in via CRISPR-mediated HDR can be achieved with readily synthesised single stranded DNAs of ~200 nucleotides that include the tag and sufficient homology arms to mediate recombination. Alternatively, the tags can be introduced with removable markers that provide convenient reporters for successful recombination and subsequent genetic crossing. These can be readily removed from the genome via one of recombination systems (e.g., Cre-*Lox* or Flp-*FRT*) described above, leaving only the tag of interest and a short *LoxP* or *FRT* site [69,135] (Figure 3e). The development of 'scarless' techniques can further limit the introduction of unwanted additional sequences into the genome [98]. A method combining both CRISPR and *piggyBac* transposase utilises the pHD-ScarlessDsRed vector, available from the *Drosophila* Genomics Resource Center (http://flycrispr.molbio.wisc.edu/scarless). Another possibility is that *LoxP* or *FRT* sites could be replaced by a gRNA to efficiently mediate the marker removal via CRISPR, however, a limitation of this approach is that NHEJ events could lead to the recovery of INDELs.

The use of CRISPR/Cas9 for tagging proteins in situ is becoming widespread and opens up genome biology approaches to virtually any desired protein in the genome: for example, ChIP or RNA immunoprecipitation (RiP) methods for identifying the binding targets of DNA or RNA binding proteins respectively, rely on the availability of highly specific antibodies. Even when good antibodies are available, it is often difficult to make accurate comparisons between samples because different antibodies do not behave in the same way. The possibility of performing such experiments for multiple different proteins with a single antibody is clearly attractive and an obvious route to achieving this is by the introduction of in situ epitope tags. Multiply RNA binding proteins have been engineered with V5 and FLAG tags [136]. Similarly, methods for introducing FLAG (CETCh-seq) [137] or FLAG and fluorescence tags (cmChIP-Seq) [138] have also been developed recently [138]. The latter combines CRISPR with microhomology mediated end joining (MMEJ) to tag transcription factors and, while it has currently been shown to be successful in human cell lines, MMEJ is certainly active in *Drosophila* [139].

Koles et al. [140] developed a strategy to enable tissue specific tagging of endogenous loci in *Drosophila* (T-STEP) by combining CRISPR-Cas9 with a yeast recombinase (Rippase). In brief, a knock-in cassette comprising tandem Rippase-specific Recognition Sequences (RRS) in frame with the targeted protein brackets a strong lethality selection marker (Golic+ [141]) to ensure high efficiency targeting: an in frame GFP is 3′ to the RRS site. Tissue specific expression of the Rippase via the Gal4-*UAS* system allows removal of the T-STEP cassette and puts the GFP in frame with the protein. The authors used this approach for live imaging of endosomal components. In cultured cells, Kunzelmann et al. [142] recently described a simplified protocol for rapid genome engineering that indicates homology arms as short as 60 bp are sufficient for efficient HDR. Their work provides a number of improvements to the system in terms of gRNA design and expression, making routine tagging of proteins in cultured cells feasible. Interesting, these authors more recently showed that the presence of selection markers in vectors for engineering cell lines induced siRNA-mediated changes in gene expression and thus removal of the marker when using such approaches [143] is desirable.

While specific reports of in vivo tagging in *Drosophila* are currently limited, anecdotally it is clear that many laboratories are actively using the technology and it is obvious that CRISPR mediated strategies will be widely applicable for tagging protein isoforms. For example, we have recently used fluorescent or small epitope tags to specifically tag different isoforms of *shaggy*, *ventral nervous system defective* and *roadkill* amongst others, using both direct tagging and Cre-*Lox* directed marker removal, with high efficiency (D. Korona and S. Russell, unpublished).

4. Conclusions

Taken together, the methods we have described here are increasing the utility of the already well-established *Drosophila* model for exploring developmental processes. From the generation of specific mutations through to the ability to track proteins in real time in vivo, genome engineering is opening new possibilities for characterising the deployment of the genome: from the characterisation of gene regulatory networks through post-transcriptional gene control, to the cell biology of proteins and protein complexes. These developments will continue to keep the fly at the forefront as an experimentally tractable system for understanding metazoan biology.

Author Contributions: All authors contributed to the background research and writing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rubin, G.M.; Lewis, E.B. A brief history of *Drosophila*'s contributions to genome research. *Science* **2000**, *287*, 2216–2218. [CrossRef] [PubMed]
2. Arias, A.M. *Drosophila* melanogaster and the development of biology in the 20th century. *Methods Mol. Biol.* **2008**, *420*, 1–25. [CrossRef] [PubMed]
3. Adryan, B.; Russell, S. Genome mapping and genomics in *Drosophila*. In *Genome Mapping and Genomics in Laboratory Animals*; Denny, P., Kole, C., Eds.; Springer: Berilin, Germany, 2012; Volume 4, pp. 31–86, ISBN 978-3-642-31315-8.
4. Russell, S. From sequence to function: The impact of the genome sequence on *Drosophila* biology. *Brief. Funct. Genom.* **2012**, *11*, 333–335. [CrossRef] [PubMed]
5. Rong, Y.S.; Golic, K.G. Gene targeting by homologous recombination in *Drosophila*. *Science* **2000**, *288*, 2013–2018. [CrossRef] [PubMed]
6. Gong, W.J.; Golic, K.G. Ends-out, or replacement, gene targeting in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2556–2561. [CrossRef] [PubMed]
7. Ryder, E.; Russell, S. Transposable elements as tools for genomics and genetics in *Drosophila*. *Brief. Funct. Genom. Proteom.* **2003**, *2*, 57–71. [CrossRef]
8. Venken, K.J.; Bellen, H.J. Chemical mutagens, transposons, and transgenes to interrogate gene function in *Drosophila* melanogaster. *Methods* **2014**, *68*, 15–28. [CrossRef] [PubMed]
9. Rubin, G.M.; Spradling, A.C. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **1982**, *218*, 348–353. [CrossRef] [PubMed]
10. Brand, A.H.; Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **1993**, *118*, 401–415. [PubMed]
11. O'Kane, C.J.; Gehring, W.J. Detection in situ of genomic regulatory elements in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 9123–9127. [CrossRef] [PubMed]
12. Wilson, C.; Pearson, R.K.; Bellen, H.J.; O'Kane, C.J.; Grossniklaus, U.; Gehring, W.J. P-element-mediated enhancer detection: An efficient method for isolating and characterizing developmentally regulated genes in *Drosophila*. *Genes. Dev.* **1989**, *3*, 1301–1313. [CrossRef] [PubMed]
13. Rørth, P. A modular misexpression screen in *Drosophila* detecting tissue-specific phenotypes. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12418–12422. [CrossRef] [PubMed]
14. Spradling, A.C.; Bellen, H.J.; Hoskins, R.A. *Drosophila* p elements preferentially transpose to replication origins. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 15948–15953. [CrossRef] [PubMed]
15. Loukeris, T.G.; Arcà, B.; Livadaras, I.; Dialektaki, G.; Savakis, C. Introduction of the transposable element minos into the germ line of *Drosophila* melanogaster. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9485–9489. [CrossRef] [PubMed]
16. Lobo, N.; Li, X.; Fraser, M.J. Transposition of the piggybac element in embryos of *Drosophila* melanogaster, aedes aegypti and trichoplusia ni. *Mol. Gen. Genet.* **1999**, *261*, 803–810. [CrossRef] [PubMed]

17. Groth, A.C.; Fish, M.; Nusse, R.; Calos, M.P. Construction of transgenic *Drosophila* by using the site-specific integrase from phage phic31. *Genetics* **2004**, *166*, 1775–1782. [CrossRef] [PubMed]

18. Venken, K.J.; Sarrion-Perdigones, A.; Vandeventer, P.J.; Abel, N.S.; Christiansen, A.E.; Hoffman, K.L. Genome engineering: *Drosophila* melanogaster and beyond. *Wiley Interdiscip. Rev. Dev. Biol.* **2016**, *5*, 233–267. [CrossRef] [PubMed]

19. Golic, K.G.; Lindquist, S. The flp recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome. *Cell* **1989**, *59*, 499–509. [CrossRef]

20. Nakazawa, N.; Taniguchi, K.; Okumura, T.; Maeda, R.; Matsuno, K. A novel cre/loxp system for mosaic gene expression in the *Drosophila* embryo. *Dev. Dyn.* **2012**, *241*, 965–974. [CrossRef] [PubMed]

21. Del Valle Rodríguez, A.; Didiano, D.; Desplan, C. Power tools for gene expression and clonal analysis in *Drosophila*. *Nat. Methods* **2011**, *9*, 47–55. [CrossRef] [PubMed]

22. Potter, C.J.; Luo, L. Using the q system in *Drosophila* melanogaster. *Nat. Protoc.* **2011**, *6*, 1105–1120. [CrossRef] [PubMed]

23. Boulina, M.; Samarajeewa, H.; Baker, J.D.; Kim, M.D.; Chiba, A. Live imaging of multicolor-labeled cells in *Drosophila*. *Development* **2013**, *140*, 1605–1613. [CrossRef] [PubMed]

24. Griffin, R.; Binari, R.; Perrimon, N. Genetic odyssey to generate marked clones in *Drosophila* mosaics. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 4756–4763. [CrossRef] [PubMed]

25. Pantazis, P.; Supatto, W. Advances in whole-embryo imaging: A quantitative transition is underway. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 327–339. [CrossRef] [PubMed]

26. Owald, D.; Lin, S.; Waddell, S. Light, heat, action: Neural control of fruit fly behaviour. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **2015**, *370*, 20140211. [CrossRef] [PubMed]

27. Viktorin, G. Using marcm to study *Drosophila* brain development. *Methods Mol. Biol.* **2014**, *1082*, 79–96. [CrossRef] [PubMed]

28. Parks, A.L.; Cook, K.R.; Belvin, M.; Dompe, N.A.; Fawcett, R.; Huppert, K.; Tan, L.R.; Winter, C.G.; Bogart, K.P.; Deal, J.E.; et al. Systematic generation of high-resolution deletion coverage of the *Drosophila* melanogaster genome. *Nat. Genet.* **2004**, *36*, 288–292. [CrossRef] [PubMed]

29. Ryder, E.; Ashburner, M.; Bautista-Llacer, R.; Drummond, J.; Webster, J.; Johnson, G.; Morley, T.; Chan, Y.S.; Blows, F.; Coulson, D.; et al. The drosdel deletion collection: A *Drosophila* genomewide chromosomal deficiency resource. *Genetics* **2007**, *177*, 615–629. [CrossRef] [PubMed]

30. Golic, K.G.; Golic, M.M. Engineering the *Drosophila* genome: Chromosome rearrangements by design. *Genetics* **1996**, *144*, 1693–1711. [PubMed]

31. Bieli, D.; Kanca, O.; Gohl, D.; Denes, A.; Schedl, P.; Affolter, M.; Müller, M. The *Drosophila* melanogaster mutants apblot and apxasta affect an essential apterous wing enhancer. *G3 Genes Genomes Genet.* **2015**, *5*, 1129–1143. [CrossRef] [PubMed]

32. Lee, H.; Cho, D.Y.; Whitworth, C.; Eisman, R.; Phelps, M.; Roote, J.; Kaufman, T.; Cook, K.; Russell, S.; Przytycka, T.; et al. Effects of gene dose, chromatin, and network topology on expression in *Drosophila* melanogaster. *PLoS Genet.* **2016**, *12*, e1006295. [CrossRef] [PubMed]

33. Meadows, L.A.; Chan, Y.S.; Roote, J.; Russell, S. Neighbourhood continuity is not required for correct testis gene expression in *Drosophila*. *PLoS Biol.* **2010**, *8*, e1000552. [CrossRef] [PubMed]

34. Cook, R.K.; Christensen, S.J.; Deal, J.A.; Coburn, R.A.; Deal, M.E.; Gresens, J.M.; Kaufman, T.C.; Cook, K.R. The generation of chromosomal deletions to provide extensive coverage and subdivision of the *Drosophila* melanogaster genome. *Genome Biol.* **2012**, *13*, R21. [CrossRef] [PubMed]

35. Bischof, J.; Maeda, R.K.; Hediger, M.; Karch, F.; Basler, K. An optimized transgenesis system for *Drosophila* using germ-line-specific phic31 integrases. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3312–3317. [CrossRef] [PubMed]

36. Markstein, M.; Pitsouli, C.; Villalta, C.; Celniker, S.E.; Perrimon, N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nat. Genet.* **2008**, *40*, 476–483. [CrossRef] [PubMed]

37. Knapp, J.M.; Chung, P.; Simpson, J.H. Generating customized transgene landing sites and multi-transgene arrays in *Drosophila* using phic31 integrase. *Genetics* **2015**, *199*, 919–934. [CrossRef] [PubMed]

38. Bateman, J.R.; Lee, A.M.; Wu, C.T. Site-specific transformation of *Drosophila* via phic31 integrase-mediated cassette exchange. *Genetics* **2006**, *173*, 769–777. [CrossRef] [PubMed]

39. Venken, K.J.; He, Y.; Hoskins, R.A.; Bellen, H.J. P[acman]: A bac transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* **2006**, *314*, 1747–1751. [CrossRef] [PubMed]

40. Ejsmont, R.K.; Sarov, M.; Winkler, S.; Lipinski, K.A.; Tomancak, P. A toolkit for high-throughput, cross-species gene engineering in *Drosophila*. *Nat. Methods* **2009**, *6*, 435–437. [CrossRef] [PubMed]

41. Sarov, M.; Barz, C.; Jambor, H.; Hein, M.Y.; Schmied, C.; Suchold, D.; Stender, B.; Janosch, S.; Vikas, V.K.J.; Krishnan, R.T.; et al. A genome-wide resource for the analysis of protein localisation in *Drosophila*. *Elife* **2016**, *5*, e12068. [CrossRef] [PubMed]

42. Morin, X.; Daneman, R.; Zavortink, M.; Chia, W. A protein trap strategy to detect gfp-tagged proteins expressed from their endogenous loci in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 15050–15055. [CrossRef] [PubMed]

43. Aleksic, J.; Lazic, R.; Müller, I.; Russell, S.R.; Adryan, B. Biases in *Drosophila* melanogaster protein trap screens. *BMC Genom.* **2009**, *10*, 249. [CrossRef] [PubMed]

44. Buszczak, M.; Paterno, S.; Lighthouse, D.; Bachman, J.; Planck, J.; Owen, S.; Skora, A.D.; Nystul, T.G.; Ohlstein, B.; Allen, A.; et al. The carnegie protein trap library: A versatile tool for *Drosophila* developmental studies. *Genetics* **2007**, *175*, 1505–1531. [CrossRef] [PubMed]

45. Quiñones-Coello, A.T.; Petrella, L.N.; Ayers, K.; Melillo, A.; Mazzalupo, S.; Hudson, A.M.; Wang, S.; Castiblanco, C.; Buszczak, M.; Hoskins, R.A.; et al. Exploring strategies for protein trapping in *Drosophila*. *Genetics* **2007**, *175*, 1089–1104. [CrossRef] [PubMed]

46. Lowe, N.; Rees, J.S.; Roote, J.; Ryder, E.; Armean, I.M.; Johnson, G.; Drummond, E.; Spriggs, H.; Drummond, J.; Magbanua, J.P.; et al. Analysis of the expression patterns, subcellular localisations and interaction partners of *Drosophila* proteins using a pigp protein trap library. *Development* **2014**, *141*, 3994–4005. [CrossRef] [PubMed]

47. Lye, C.M.; Naylor, H.W.; Sanson, B. Subcellular localisations of the cpti collection of yfp-tagged proteins in *Drosophila* embryos. *Development* **2014**, *141*, 4006–4017. [CrossRef] [PubMed]

48. Rees, J.S.; Lowe, N.; Armean, I.M.; Roote, J.; Johnson, G.; Drummond, E.; Spriggs, H.; Ryder, E.; Russell, S.; St Johnston, D.; et al. In vivo analysis of proteomes and interactomes using parallel affinity capture (ipac) coupled to mass spectrometry. *Mol. Cell. Proteom.* **2011**, *10*. [CrossRef] [PubMed]

49. Nashchekin, D.; Fernandes, A.R.; St Johnston, D. Patronin/shot cortical foci assemble the noncentrosomal microtubule array that specifies the *Drosophila* anterior-posterior axis. *Dev. Cell* **2016**, *38*, 61–72. [CrossRef] [PubMed]

50. Byri, S.; Misra, T.; Syed, Z.A.; Bätz, T.; Shah, J.; Boril, L.; Glashauser, J.; Aegerter-Wilmsen, T.; Matzat, T.; Moussian, B.; et al. The triple-repeat protein anakonda controls epithelial tricellular junction formation in *Drosophila*. *Dev. Cell* **2015**, *33*, 535–548. [CrossRef] [PubMed]

51. Norman, M.; Vuilleumier, R.; Springhorn, A.; Gawlik, J.; Pyrowolakis, G. Pentagone internalises glypicans to fine-tune multiple signalling pathways. *Elife* **2016**, *5*. [CrossRef] [PubMed]

52. Choo, S.W.; White, R.; Russell, S. Genome-wide analysis of the binding of the hox protein ultrabithorax and the hox cofactor homothorax in *Drosophila*. *PLoS ONE* **2011**, *6*, e14778. [CrossRef] [PubMed]

53. Venken, K.J.; Schulze, K.L.; Haelterman, N.A.; Pan, H.; He, Y.; Evans-Holm, M.; Carlson, J.W.; Levis, R.W.; Spradling, A.C.; Hoskins, R.A.; et al. Mimic: A highly versatile transposon insertion resource for engineering *Drosophila* melanogaster genes. *Nat. Methods* **2011**, *8*, 737–743. [CrossRef] [PubMed]

54. Nagarkar-Jaiswal, S.; DeLuca, S.Z.; Lee, P.T.; Lin, W.W.; Pan, H.; Zuo, Z.; Lv, J.; Spradling, A.C.; Bellen, H.J. A genetic toolkit for tagging intronic mimic containing genes. *Elife* **2015**, *4*. [CrossRef] [PubMed]

55. Nagarkar-Jaiswal, S.; Lee, P.T.; Campbell, M.E.; Chen, K.; Anguiano-Zarate, S.; Gutierrez, M.C.; Busby, T.; Lin, W.W.; He, Y.; Schulze, K.L.; et al. A library of mimics allows tagging of genes and reversible, spatial and temporal knockdown of proteins in *Drosophila*. *Elife* **2015**, *4*. [CrossRef] [PubMed]

56. Caussinus, E.; Kanca, O.; Affolter, M. Fluorescent fusion protein knockout mediated by anti-gfp nanobody. *Nat. Struct. Mol. Biol.* **2011**, *19*, 117–121. [CrossRef] [PubMed]

57. Neumüller, R.A.; Wirtz-Peitz, F.; Lee, S.; Kwon, Y.; Buckner, M.; Hoskins, R.A.; Venken, K.J.; Bellen, H.J.; Mohr, S.E.; Perrimon, N. Stringent analysis of gene function and protein-protein interactions using fluorescently tagged genes. *Genetics* **2012**, *190*, 931–940. [CrossRef] [PubMed]

58. Trost, M.; Blattner, A.C.; Lehner, C.F. Regulated protein depletion by the auxin-inducible degradation system in *Drosophila* melanogaster. *Fly* **2016**, *10*, 35–46. [CrossRef] [PubMed]

59. Hsu, P.D.; Lander, E.S.; Zhang, F. Development and applications of crispr-cas9 for genome engineering. *Cell* **2014**, *157*, 1262–1278. [CrossRef] [PubMed]

60. Beumer, K.J.; Carroll, D. Targeted genome engineering techniques in *Drosophila*. *Methods* **2014**, *68*, 29–37. [CrossRef] [PubMed]

61. Yu, Z.; Chen, H.; Liu, J.; Zhang, H.; Yan, Y.; Zhu, N.; Guo, Y.; Yang, B.; Chang, Y.; Dai, F.; et al. Various applications of talen- and crispr/cas9-mediated homologous recombination to modify the *Drosophila* genome. *Biol. Open* **2014**, *3*, 271–280. [CrossRef] [PubMed]

62. Cebrian-Serrano, A.; Davies, B. Crispr-cas orthologues and variants: Optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm. Genome* **2017**. [CrossRef] [PubMed]

63. Gasiunas, G.; Barrangou, R.; Horvath, P.; Siksnys, V. Cas9-crrna ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2579–E2586. [CrossRef] [PubMed]

64. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-rna-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821. [CrossRef] [PubMed]

65. Bétermier, M.; Bertrand, P.; Lopez, B.S. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genet.* **2014**, *10*, e1004086. [CrossRef] [PubMed]

66. Greene, E.C. DNA sequence alignment during homologous recombination. *J. Biol. Chem.* **2016**, *291*, 11572–11580. [CrossRef] [PubMed]

67. Greene, E.C. On the influence of protein-DNA register during homologous recombination. *Cell Cycle* **2016**, *15*, 172–175. [CrossRef] [PubMed]

68. Gratz, S.J.; Rubinstein, C.D.; Harrison, M.M.; Wildonger, J.; O'Connor-Giles, K.M. Crispr-cas9 genome editing in *Drosophila*. *Curr. Protoc. Mol. Biol.* **2015**, *111*. [CrossRef]

69. Gratz, S.J.; Ukken, F.P.; Rubinstein, C.D.; Thiede, G.; Donohue, L.K.; Cummings, A.M.; O'Connor-Giles, K.M. Highly specific and efficient crispr/cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics* **2014**, *196*, 961–971. [CrossRef] [PubMed]

70. Port, F.; Chen, H.M.; Lee, T.; Bullock, S.L. Optimized crispr/cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E2967–E2976. [CrossRef] [PubMed]

71. Schaefer, K.A.; Wu, W.H.; Colgan, D.F.; Tsang, S.H.; Bassuk, A.G.; Mahajan, V.B. Unexpected mutations after crispr-cas9 editing in vivo. *Nat. Methods* **2017**, *14*, 547–548. [CrossRef] [PubMed]

72. Lee, C.M.; Cradick, T.J.; Fine, E.J.; Bao, G. Nuclease target site selection for maximizing on-target activity and minimizing off-target effects in genome editing. *Mol. Ther.* **2016**, *24*, 475–487. [CrossRef] [PubMed]

73. Stella, S.; Montoya, G. The genome editing revolution: A crispr-cas tale off-target story. *Bioessays* **2016**, *38* (Suppl. 1), S4–S13. [CrossRef] [PubMed]

74. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A.; et al. Multiplex genome engineering using crispr/cas systems. *Science* **2013**, *339*, 819–823. [CrossRef] [PubMed]

75. Mali, P.; Aach, J.; Stranges, P.B.; Esvelt, K.M.; Moosburner, M.; Kosuri, S.; Yang, L.; Church, G.M. Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **2013**, *31*, 833–838. [CrossRef] [PubMed]

76. 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.; et al. Double nicking by rna-guided crispr cas9 for enhanced genome editing specificity. *Cell* **2013**, *154*, 1380–1389. [CrossRef] [PubMed]

77. Overballe-Petersen, S.; Harms, K.; Orlando, L.A.; Mayar, J.V.; Rasmussen, S.; Dahl, T.W.; Rosing, M.T.; Poole, A.M.; Sicheritz-Ponten, T.; Brunak, S.; et al. Bacterial natural transformation by highly fragmented and damaged DNA. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19860–19865. [CrossRef] [PubMed]

78. Cho, S.W.; Kim, S.; Kim, Y.; Kweon, J.; Kim, H.S.; Bae, S.; Kim, J.S. Analysis of off-target effects of crispr/cas-derived rna-guided endonucleases and nickases. *Genome Res.* **2014**, *24*, 132–141. [CrossRef] [PubMed]

79. Miyaoka, Y.; Berman, J.R.; Cooper, S.B.; Mayerl, S.J.; Chan, A.H.; Zhang, B.; Karlin-Neumann, G.A.; Conklin, B.R. Systematic quantification of hdr and nhej reveals effects of locus, nuclease, and cell type on genome-editing. *Sci. Rep.* **2016**, *6*, 23549. [CrossRef] [PubMed]

80. Doudna, J.A.; Charpentier, E. Genome editing. The new frontier of genome engineering with crispr-cas9. *Science* **2014**, *346*, 1258096. [CrossRef] [PubMed]

81. Guilinger, J.P.; Thompson, D.B.; Liu, D.R. Fusion of catalytically inactive cas9 to foki nuclease improves the specificity of genome modification. *Nat. Biotechnol.* **2014**, *32*, 577–582. [CrossRef] [PubMed]

82. Kleinstiver, B.P.; Pattanayak, V.; Prew, M.S.; Tsai, S.Q.; Nguyen, N.T.; Zheng, Z.; Joung, J.K. High-fidelity crispr-cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **2016**, *529*, 490–495. [CrossRef] [PubMed]

83. Slaymaker, I.M.; Gao, L.; Zetsche, B.; Scott, D.A.; Yan, W.X.; Zhang, F. Rationally engineered cas9 nucleases with improved specificity. *Science* **2016**, *351*, 84–88. [CrossRef] [PubMed]

84. Kleinstiver, B.P.; Prew, M.S.; Tsai, S.Q.; Topkar, V.V.; Nguyen, N.T.; Zheng, Z.; Gonzales, A.P.; Li, Z.; Peterson, R.T.; Yeh, J.R.; et al. Engineered crispr-cas9 nucleases with altered pam specificities. *Nature* **2015**, *523*, 481–485. [CrossRef] [PubMed]

85. Zetsche, B.; Gootenberg, J.S.; Abudayyeh, O.O.; Slaymaker, I.M.; Makarova, K.S.; Essletzbichler, P.; Volz, S.E.; Joung, J.; van der Oost, J.; Regev, A.; et al. Cpf1 is a single rna-guided endonuclease of a class 2 crispr-cas system. *Cell* **2015**, *163*, 759–771. [CrossRef] [PubMed]

86. Shmakov, S.; Abudayyeh, O.O.; Makarova, K.S.; Wolf, Y.I.; Gootenberg, J.S.; Semenova, E.; Minakhin, L.; Joung, J.; Konermann, S.; Severinov, K.; et al. Discovery and functional characterization of diverse class 2 crispr-cas systems. *Mol. Cell.* **2015**, *60*, 385–397. [CrossRef] [PubMed]

87. 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. Genome engineering of *Drosophila* with the crispr rna-guided cas9 nuclease. *Genetics* **2013**, *194*, 1029–1035. [CrossRef] [PubMed]

88. Hilton, I.B.; Gersbach, C.A. Enabling functional genomics with genome engineering. *Genome Res.* **2015**, *25*, 1442–1455. [CrossRef] [PubMed]

89. Essletzbichler, P.; Konopka, T.; Santoro, F.; Chen, D.; Gapp, B.V.; Kralovics, R.; Brummelkamp, T.R.; Nijman, S.M.; Burckstummer, T. Megabase-scale deletion using crispr/cas9 to generate a fully haploid human cell line. *Genome Res.* **2014**, *24*, 2059–2065. [CrossRef] [PubMed]

90. Auer, T.O.; Duroure, K.; De Cian, A.; Concordet, J.P.; Del Bene, F. Highly efficient crispr/cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res.* **2014**, *24*, 142–153. [CrossRef] [PubMed]

91. He, X.; Tan, C.; Wang, F.; Wang, Y.; Zhou, R.; Cui, D.; You, W.; Zhao, H.; Ren, J.; Feng, B. Knock-in of large reporter genes in human cells via crispr/cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res.* **2016**, *44*, e85. [CrossRef] [PubMed]

92. Böttcher, R.; Hollmann, M.; Merk, K.; Nitschko, V.; Obermaier, C.; Philippou-Massier, J.; Wieland, I.; Gaul, U.; Förstemann, K. Efficient chromosomal gene modification with crispr/cas9 and pcr-based homologous recombination donors in cultured *Drosophila* cells. *Nucleic Acids Res.* **2014**, *42*, e89. [CrossRef] [PubMed]

93. Housden, B.E.; Lin, S.; Perrimon, N. Cas9-based genome editing in *Drosophila*. *Methods Enzymol.* **2014**, *546*, 415–439. [CrossRef] [PubMed]

94. Salsman, J.; Dellaire, G. Precision genome editing in the crispr era. *Biochem. Cell Biol.* **2017**, *95*, 187–201. [CrossRef] [PubMed]

95. Yang, L.; Guell, M.; Byrne, S.; Yang, J.L.; De Los Angeles, A.; Mali, P.; Aach, J.; Kim-Kiselak, C.; Briggs, A.W.; Rios, X.; et al. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* **2013**, *41*, 9049–9061. [CrossRef] [PubMed]

96. Komor, A.C.; Badran, A.H.; Liu, D.R. Crispr-based technologies for the manipulation of eukaryotic genomes. *Cell* **2017**, *169*, 559. [CrossRef] [PubMed]

97. Gaudelli, N.M.; Komor, A.C.; Rees, H.A.; Packer, M.S.; Badran, A.H.; Bryson, D.I. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **2017**. [CrossRef] [PubMed]

98. Lamb, A.M.; Walker, E.A.; Wittkopp, P.J. Tools and strategies for scarless allele replacement in *Drosophila* using crispr/cas9. *Fly* **2017**, *11*, 53–64. [CrossRef] [PubMed]

99. Byrne, S.M.; Ortiz, L.; Mali, P.; Aach, J.; Church, G.M. Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Res.* **2015**, *43*, e21. [CrossRef] [PubMed]

100. Zheng, Q.; Cai, X.; Tan, M.H.; Schaffert, S.; Arnold, C.P.; Gong, X.; Chen, C.Z.; Huang, S. Precise gene deletion and replacement using the crispr/cas9 system in human cells. *Biotechniques* **2014**, *57*, 115–124. [CrossRef] [PubMed]

101. Zhao, Y.; Zhang, C.; Liu, W.; Gao, W.; Liu, C.; Song, G.; Li, W.X.; Mao, L.; Chen, B.; Xu, Y.; et al. An alternative strategy for targeted gene replacement in plants using a dual-sgrna/cas9 design. *Sci. Rep.* **2016**, *6*, 23890. [CrossRef] [PubMed]

102. Adikusuma, F.; Pederick, D.; McAninch, D.; Hughes, J.; Thomas, P. Functional equivalence of the sox2 and sox3 transcription factors in the developing mouse brain and testes. *Genetics* **2017**, *206*, 1495–1503. [CrossRef] [PubMed]

103. Evans, T.A. Crispr-based gene replacement reveals evolutionarily conserved axon guidance functions of *Drosophila* robo3 and tribolium robo2/3. *Evodevo* **2017**, *8*, 10. [CrossRef] [PubMed]

104. Spitzweck, B.; Brankatschk, M.; Dickson, B.J. Distinct protein domains and expression patterns confer divergent axon guidance functions for *Drosophila* robo receptors. *Cell* **2010**, *140*, 409–420. [CrossRef] [PubMed]

105. Ewen-Campen, B.; Mohr, S.E.; Hu, Y.; Perrimon, N. Accessing the phenotype gap: Enabling systematic investigation of paralog functional complexity with crispr. *Dev. Cell* **2017**, *43*, 6–9. [CrossRef] [PubMed]

106. Overton, P.M.; Meadows, L.A.; Urban, J.; Russell, S. Evidence for differential and redundant function of the sox genes dichaete and soxn during cns development in *Drosophila*. *Development* **2002**, *129*, 4219–4228. [PubMed]

107. Sánchez-Soriano, N.; Russell, S. Regulatory mutations of the *Drosophila* sox gene dichaete reveal new functions in embryonic brain and hindgut development. *Dev. Biol.* **2000**, *220*, 307–321. [CrossRef] [PubMed]

108. Ren, X.; Yang, Z.; Xu, J.; Sun, J.; Mao, D.; Hu, Y.; Yang, S.J.; Qiao, H.H.; Wang, X.; Hu, Q.; et al. Enhanced specificity and efficiency of the crispr/cas9 system with optimized sgrna parameters in *Drosophila*. *Cell Rep.* **2014**, *9*, 1151–1162. [CrossRef] [PubMed]

109. Port, F.; Bullock, S.L. Augmenting crispr applications in *Drosophila* with trna-flanked sgrnas. *Nat. Methods* **2016**, *13*, 852–854. [CrossRef] [PubMed]

110. Yang, H.; Wang, H.; Shivalila, C.S.; Cheng, A.W.; Shi, L.; Jaenisch, R. One-step generation of mice carrying reporter and conditional alleles by crispr/cas-mediated genome engineering. *Cell* **2013**, *154*, 1370–1379. [CrossRef] [PubMed]

111. Arribere, J.A.; Bell, R.T.; Fu, B.X.; Artiles, K.L.; Hartman, P.S.; Fire, A.Z. Efficient marker-free recovery of custom genetic modifications with crispr/cas9 in caenorhabditis elegans. *Genetics* **2014**, *198*, 837–846. [CrossRef] [PubMed]

112. Ward, J.D. Rapid and precise engineering of the caenorhabditis elegans genome with lethal mutation co-conversion and inactivation of nhej repair. *Genetics* **2015**, *199*, 363–377. [CrossRef] [PubMed]

113. Kane, N.S.; Vora, M.; Varre, K.J.; Padgett, R.W. Efficient screening of crispr/cas9-induced events in *Drosophila* using a co-crispr strategy. *G3 Genes Genomes Genet.* **2017**, *7*, 87–93. [CrossRef] [PubMed]

114. Ge, D.T.; Tipping, C.; Brodsky, M.H.; Zamore, P.D. Rapid screening for crispr-directed editing of the *Drosophila* genome using white coconversion. *G3 Genes Genomes Genet.* **2016**, *6*, 3197–3206. [CrossRef]

115. Kondo, S.; Ueda, R. Highly improved gene targeting by germline-specific cas9 expression in *Drosophila*. *Genetics* **2013**, *195*, 715–721. [CrossRef] [PubMed]

116. Ren, X.; Sun, J.; Housden, B.E.; Hu, Y.; Roesel, C.; Lin, S.; Liu, L.P.; Yang, Z.; Mao, D.; Sun, L.; et al. Optimized gene editing technology for *Drosophila* melanogaster using germ line-specific cas9. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19012–19017. [CrossRef] [PubMed]

117. Xue, Z.; Wu, M.; Wen, K.; Ren, M.; Long, L.; Zhang, X.; Gao, G. Crispr/cas9 mediates efficient conditional mutagenesis in *Drosophila*. *G3 Genes Genomes Genet.* **2014**, *4*, 2167–2173. [CrossRef] [PubMed]

118. Zhang, X.; Koolhaas, W.H.; Schnorrer, F. A versatile two-step crispr- and rmce-based strategy for efficient genome engineering in *Drosophila*. *G3 Genes Genomes Genet.* **2014**, *4*, 2409–2418. [CrossRef] [PubMed]

119. Stern, D.L.; Frankel, N. The structure and evolution of cis-regulatory regions: The shavenbaby story. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **2013**, *368*, 20130028. [CrossRef] [PubMed]

120. Wilczynski, B.; Furlong, E.E. Challenges for modeling global gene regulatory networks during development: Insights from *Drosophila*. *Dev. Biol.* **2010**, *340*, 161–169. [CrossRef] [PubMed]

121. Mendoza-Garcia, P.; Hugosson, F.; Fallah, M.; Higgins, M.L.; Iwasaki, Y.; Pfeifer, K.; Wolfstetter, G.; Varshney, G.; Popichenko, D.; Gergen, J.P.; et al. The zic family homologue odd-paired regulates alk expression in *Drosophila*. *PLoS Genet.* **2017**, *13*, e1006617. [CrossRef] [PubMed]

122. Zandvakili, A.; Gebelein, B. Mechanisms of specificity for hox factor activity. *J. Dev. Biol.* **2016**, *4*, 16. [CrossRef] [PubMed]

123. Rogers, W.A.; Goyal, Y.; Yamaya, K.; Shvartsman, S.Y.; Levine, M.S. Uncoupling neurogenic gene networks in the *Drosophila* embryo. *Genes Dev.* **2017**, *31*, 634–638. [CrossRef] [PubMed]

124. Lagha, M.; Bothma, J.P.; Levine, M. Mechanisms of transcriptional precision in animal development. *Trends Genet.* **2012**, *28*, 409–416. [CrossRef] [PubMed]

125. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74. [CrossRef]

126. Lee, Y.; Rio, D.C. Mechanisms and regulation of alternative pre-mrna splicing. *Annu. Rev. Biochem.* **2015**, *84*, 291–323. [CrossRef] [PubMed]

127. Beagan, J.A.; Phillips-Cremins, J.E. Crispr/cas9 genome editing throws descriptive 3-D genome folding studies for a loop. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2016**, *8*, 286–299. [CrossRef] [PubMed]

128. Mir, M.; Reimer, A.; Haines, J.E.; Li, X.Y.; Stadler, M.; Garcia, H.; Eisen, M.B.; Darzacq, X. Dense bicoid hubs accentuate binding along the morphogen gradient. *Genes Dev.* **2017**, *31*, 1784–1794. [CrossRef] [PubMed]

129. Crews, S. Creating cell type-specific mutants by enhancer mutagenesis. *Genes Dev.* **2017**, *31*, 629–631. [CrossRef] [PubMed]

130. Wei, Y.; Zhang, S.; Shang, S.; Zhang, B.; Li, S.; Wang, X.; Wang, F.; Su, J.; Wu, Q.; Liu, H.; et al. Sea: A super-enhancer archive. *Nucleic Acids Res.* **2016**, *44*, D172–D179. [CrossRef] [PubMed]

131. Thorn, K. Genetically encoded fluorescent tags. *Mol. Biol. Cell* **2017**, *28*, 848–857. [CrossRef] [PubMed]

132. Rees, J.S.; Lilley, K.S.; Jackson, A.P. Silac-ipac: A quantitative method for distinguishing genuine from non-specific components of protein complexes by parallel affinity capture. *J. Proteom.* **2015**, *115*, 143–156. [CrossRef] [PubMed]

133. Cabantous, S.; Terwilliger, T.C.; Waldo, G.S. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat. Biotechnol.* **2005**, *23*, 102–107. [CrossRef] [PubMed]

134. Kamiyama, D.; Sekine, S.; Barsi-Rhyne, B.; Hu, J.; Chen, B.; Gilbert, L.A.; Ishikawa, H.; Leonetti, M.D.; Marshall, W.F.; Weissman, J.S.; et al. Versatile protein tagging in cells with split fluorescent protein. *Nat. Commun.* **2016**, *7*, 11046. [CrossRef] [PubMed]

135. Li, Q.; Barish, S.; Okuwa, S.; Volkan, P.C. Examination of endogenous rotund expression and function in developing *Drosophila* olfactory system using crispr-cas9-mediated protein tagging. *G3 Genes Genomes Genet.* **2015**, *5*, 2809–2816. [CrossRef] [PubMed]

136. Van Nostrand, E.L.; Gelboin-Burkhart, C.; Wang, R.; Pratt, G.A.; Blue, S.M.; Yeo, G.W. Crispr/cas9-mediated integration enables tag-eclip of endogenously tagged rna binding proteins. *Methods* **2017**, *118–119*, 50–59. [CrossRef] [PubMed]

137. Savic, D.; Partridge, E.C.; Newberry, K.M.; Smith, S.B.; Meadows, S.K.; Roberts, B.S.; Mackiewicz, M.; Mendenhall, E.M.; Myers, R.M. Cetch-seq: Crispr epitope tagging chip-seq of DNA-binding proteins. *Genome Res.* **2015**, *25*, 1581–1589. [CrossRef] [PubMed]

138. Xiong, X.; Zhang, Y.; Yan, J.; Jain, S.; Chee, S.; Ren, B.; Zhao, H. A scalable epitope tagging approach for high throughput chip-seq analysis. *ACS Synth. Biol.* **2017**, *6*, 1034–1042. [CrossRef] [PubMed]

139. Yu, A.M.; McVey, M. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res.* **2010**, *38*, 5706–5717. [CrossRef] [PubMed]

140. Koles, K.; Yeh, A.R.; Rodal, A.A. Tissue-specific tagging of endogenous loci in *Drosophila* melanogaster. *Biol. Open* **2015**, *5*, 83–89. [CrossRef] [PubMed]

141. Chen, H.M.; Huang, Y.; Pfeiffer, B.D.; Yao, X.; Lee, T. An enhanced gene targeting toolkit for *Drosophila*: Golic+. *Genetics* **2015**, *199*, 683–694. [CrossRef] [PubMed]

142. Kunzelmann, S.; Böttcher, R.; Schmidts, I.; Förstemann, K. A comprehensive toolbox for genome editing in cultured *Drosophila* melanogaster cells. *G3 Genes Genomes Genet.* **2016**, *6*, 1777–1785. [CrossRef] [PubMed]

143. Kunzelmann, S.; Förstemann, K. Reversible perturbations of gene regulation after genome editing in *Drosophila* cells. *PLoS ONE* **2017**, *12*, e0180135. [CrossRef] [PubMed]