© 2013. Published by The Company of Biologists Ltd | Development (2013) 140, 1-? doi:10.1242/dev.100933

RESEARCH ARTICLE

TECHNIQUES AND RESOURCES

Accelerated homologous recombination and subsequent genome modification in *Drosophila*

Luis Alberto Baena-Lopez*, Cyrille Alexandre*, Alice Mitchell, Laurynas Pasakarnis and Jean-Paul Vincent‡

ABSTRACT

Gene targeting by 'ends-out' homologous recombination enables the deletion of genomic sequences and concurrent introduction of exogenous DNA with base-pair precision without sequence constraint. In *Drosophila*, this powerful technique has remained laborious and hence seldom implemented. We describe a targeting vector and protocols that achieve this at high frequency and with very few false positives in *Drosophila*, either with a two-generation crossing scheme or by direct injection in embryos. The frequency of injection-mediated gene targeting can be further increased with CRISPR-induced double-strand breaks within the region to be deleted, thus making homologous recombination almost as easy as conventional transgenesis. Our targeting vector replaces genomic sequences with a multifunctional fragment comprising an easy-to-select genetic marker, a fluorescent reporter, as well as an *attP* site, which acts as a landing platform for reintegration vectors. These vectors allow the insertion of a variety of transcription reporters or cDNAs to express tagged or mutant isoforms at endogenous levels. In addition, they pave the way for difficult experiments such as tissue-specific allele switching and functional analysis in post-mitotic or polyploid cells. Therefore, our method retains the advantages of homologous recombination while capitalising on the mutagenic power of CRISPR.

KEY WORDS: *Drosophila*, Functional genomics, Gene targeting, Homologous recombination

INTRODUCTION

In mice, gene targeting is achieved by homologous recombination in embryonic stem (ES) cells, which are amenable to straightforward selection (Capecchi, 2005). In most other model species, no ES cells are available and, therefore, alternative approaches are needed. Golic and collaborators were the first to describe protocols for gene targeting by homologous recombination in *Drosophila* (Gong and Golic, 2003; Rong and Golic, 2000). Soon thereafter, gene targeting by 'ends out' homologous recombination was used to delete genomic fragments while at the same time introducing exogenous sequences with base-pair precision (Gong and Golic, 2003; Huang et al., 2009; Huang et al., 2008; Liu et al., 2012; Xie and Golic, 2004). Current protocols for homologous recombination involve several key steps (Gong and Golic, 2003; Huang et al., 2009; Huang et al., 2008; Weng et al., 2009; Xie and Golic, 2004; Zhou et al., 2012). First, homology arms, typically 3-5 kb, are ligated into the targeting vector, which is then inserted at a random location in the genome by P-element mediated transformation. As the targeting vector comprises *FLP* recombination targets (FRT) and *I-SceI* restriction sites, it can be released from the random genomic location and linearised to gain access to the target locus (Huang et al., 2009; Rong and Golic, 2000). The most commonly used targeting vector, pGX (Huang et al., 2009), comprises two key features between the homology arms (the region that becomes inserted at the locus): *mini-white*, which serves as a genetic marker; and an *attP* site to enable PhiC31-mediated insertion of exogenous sequences in the targeted locus (Bateman et al., 2006; Bischof et al., 2007; Groth et al., 2004; Venken et al., 2006). The *mini-white* genetic marker is flanked by *LoxP* sites so that it can be removed, leaving only the *attP* site (and one *LoxP* site) at the modified locus. Outside the homology arm, on the 3′ end, pGX contains a DNA fragment that expresses Reaper, a pro-apoptotic protein, when Gal4 is present. This part of the vector should not integrate in the targeted locus following accurate homologous recombination and therefore allows the negative selection of illegitimate events by crossing to a suitable Gal4 driver. So far the promises of this feature have been only partially realised because of the unavailability of Gal4 drivers that fulfil the dual conditions of being a strong activator in essential tissues while not being marked with *mini-white* so as not to interfere with the *mini-white* of the targeting vector. Moreover, the frequency of successful targeting events has remained frustratingly low. Typically, 20,000 to 100,000 flies must be screened to identify one successful targeting event (Huang et al., 2009; Huang et al., 2008; Zhou et al., 2012). Despite the palpable benefits of homologous recombination, particularly the opportunity to insert an *attP* site at any genomic location, the low frequency of recombination and the large number of false positives have deterred many laboratories from initiating homologous recombination projects.

Recently, the advent of sequence-specific endonucleases such as transcription activator-like effector nucleases (TALENs) and the RNA-guided DNA endonuclease Cas9 has dramatically opened opportunities for creating double-strand breaks (DSBs) in the genomes of yeast, flies, zebrafish and cultured mammalian cells (Bassett et al., 2013; Bedell et al., 2012; Chang et al., 2013; Cho et al., 2013; Cong et al., 2013; DiCarlo et al., 2013; Gaj et al., 2013; Gratz et al., 2013; Hwang et al., 2013; Liu et al., 2012; Mali et al., 2013). As DSB-induced non-homologous end joining (NHEJ) creates deletions, TALENs and CRISPR are effective site-specific mutagenic tools. However, these deletions are variable in length and not amenable to subsequent genome engineering. This can be overcome with concurrent introduction of homologous sequences so that the DSB is repaired by homologous recombination while exogenous sequences are inserted at the targeted locus. Proof-of-principle that this can be achieved in *Drosophila* has recently been reported (Gratz et al., 2013). DSBs and homologous recombination

MRC National Institute for Medical Research, London NW7 1AA, UK.
*These authors contributed equally to this work

‡Author for correspondence (jvincen@nimr.mrc.ac.uk)

Received 8 July 2013; Accepted 15 September 2013

Development (2013) doi:10.1242/dev.100933

were induced by co-injecting CRISPR components (guide RNA, and a Cas9-encoding plasmid), along with template DNA comprising short (50 bp) homology arms flanking an *attP* site, into *Drosophila* embryos. Successful integration of the *attP* site was demonstrated although, because of the absence of a genetic marker, successful recombinants had to be selected by PCR from pools of animals derived from injected embryos, adding a substantial amount of work to the protocol. It must be also kept in mind that CRISPR is somewhat restricted in the sequences it can target (Mali et al., 2013) and can cause off-target effects (Fu et al., 2013), a problem that could be confounded when small homology arms are used (50 bp). The results of Gratz et al. (Gratz et al., 2013) nevertheless suggest that combining CRISPR with homologous recombination is a promising approach to genome engineering. As an alternative, DSBs and homologous recombination could be induced with TALENs (Beumer et al., 2008). However, the production of plasmids encoding TALENs still requires substantial bench work, despite recent progress (Cermak et al., 2011), tempering enthusiasm for this option, at least in its current state.

We have devised vectors and protocols for homologous recombination in *Drosophila*, which are effective with and without concurrent CRISPR-induced DSBs. One CRISPR-independent approach, which requires prior integration of the targeting vector by P-element mediated, achieved a ~10-fold frequency improvement over the most recent CRISPR-independent protocol (Huang et al., 2009; Zhou et al., 2012) and reduced the number of false positives to very few. In the second approach, also CRISPR independent, the targeting vector was injected in embryos and homologous recombinants were identified in the F2 progeny with an eye colour marker screen. In the third approach, embryos were similarly injected with the targeting vector but this time with CRISPR components. Approach 1 is reliable and does not suffer with the uncertainties associated with CRISPR; it is therefore currently the safest while still achieving excellent frequencies. With the third protocol come the advantages – and current uncertainties – of CRISPR. Importantly, all three protocols use the same targeting vector, which comprises a Cre-removable reporter gene and selectable marker, as well as an *attP* site. To take full advantage of the latter, we have created a set of reintegration vectors that enable a variety of genetic manipulations including tissue-specific allele switching and the generation of genetic mosaics in post-mitotic or polyploid tissues.

RESULTS

High-frequency integration of a multifunctional DNA fragment at specific genomic locations

As a benchmark, we first tested the most commonly used *Drosophila* targeting vector (pGX) and protocol (Huang et al., 2009) with *rhogap102A* (a gene located on the 4th chromosome of *Drosophila*). The lengths of the homology arms were 5 kb (5′ arm) and 3 kb (3′ arm). From 80,000 screened flies, 365 *white*$^{+}$ flies were identified. They were crossed a second time to *221-Gal4[w-]*, with the aim of eliminating false positives missed in the first round. This left us with 170 candidates to test by PCR with primers spanning the homology arms. Of these, only one turned out to represent a true homologous recombinant. The high number of false positives is in accordance with what was previously described for other genes (Huang et al., 2008; Huang et al., 2009). To overcome the problem of false positives, we generated *ubiquitin-Gal4[3xP3-GFP]*, which was predicted to activate *UAS-reaper* throughout the larva and not only in the nervous system as *221-Gal4[w-]*. This driver was inserted in a vector that uses a fluorescent genetic marker (*3xP3-GFP*, expressed in the eye and ocelli (Horn et al., 2000). As the *3xP3-GFP* maker does not interfere with that of the targeting vector (*mini-white*), both constructs can be tracked in the same animal. None of the 364 false positives obtained during the above targeting of *rhogap102A* survived in the presence of this driver, showing its effectiveness at eliminating false positives. We next set out to improve the targeting vector, retaining important features of pGX such as *mini-white* and the *attP* site, and adding a reporter gene (Cherry) to enable rapid assessment of gene expression following gene targeting (Fig. 1A). This vector, which we call pTV^{Cherry}, was used to re-target *rhogap102A*, using the same homology arms as before. Donor flies (carrying the targeting vector at a random location) were crossed to flies expressing *FLP* and *I-SceI* (Fig. 1B) and about 200 progeny with mottled red eyes (i.e. carrying both the targeting vector and *hs-FLP, hs-I-SceI*) were collected. These were crossed to *ubiquitin-Gal4[3xP3-GFP]* to eliminate unmobilised donor chromosomes as well as false positives (Fig. 1B). Red-eyed flies resulting from this second cross were considered candidate knockouts (KOs). In this instance, two such candidates were identified from about 8000 flies examined and one was confirmed to be a homologous recombinant by PCR (Table 1; supplementary material Fig. S1). Therefore, a two-generation protocol (Fig. 1B) followed by cursory examination of about 8000 flies were sufficient to generate one KO. To confirm the effectiveness of this protocol, five more genes were targeted with the same strategy (Table 1; supplementary material Figs S1 and S2). In all cases, 200 females carrying the targeting vector and *hs-FLP, hs-I-SceI* were used and, on average, a confirmed KO was identified after screening 3000 flies (range 1000-8000). This is equivalent to a 10- to 100-fold frequency increase over previously reported results (Huang et al., 2009; Zhou et al., 2012). The number of false positives was low, approximately equal to the number of true homologous recombination events (Table 1). The small number of false positives can be directly attributed to the broad expression of *ubiquitin-Gal4[3xP3-GFP]*. The high targeting efficiency is likely due to features of pTV^{Cherry} as the same homology arms were used to target *rhogap102A* with pGX (frequency=1/80,000) and pTV^{Cherry} (frequency=1/8000). The reason for this difference is unknown but pTV^{Cherry} still appears to be 10 times more conducive to homologous recombination than pGX. This, along with an effective means of eliminating false positives, significantly facilitates homologous recombination in *Drosophila*.

Accelerated gene targeting by direct injection, with or without CRISPR

So far, homologous recombination in *Drosophila* has necessitated prior insertion of the targeting construct at a random genomic location before releasing it through genetic crosses (as described above). Targeting would be significantly accelerated if homologous recombination occurred in the germline of injected embryos. The high targeting efficiency of pTV^{Cherry} obtained with the crossing scheme suggested that it might be possible to achieve homologous recombination following injection of the targeting vector in embryos (Fig. 1C). Five different targeting plasmids were injected in the posterior region of pre-cellular embryos, along with a vector expressing *I-SceI* under the control of the germ line-specific *vasa* promoter to trigger linearization. For historical reasons, this vector (vasa-FUS) also expresses FLP and Utrophin-GFP, although these proteins are not expected to be relevant in the present protocol. In these experiments, 3000 *white*^{1118} mutant embryos were injected by Rainbow Transgenics, and the resulting adults were backcrossed to *white*^{1118} flies. The subsequent F2 progeny were screened for red eyes as this would indicate germ line transmission of the targeting

Development (2013) doi:10.1242/dev.100933

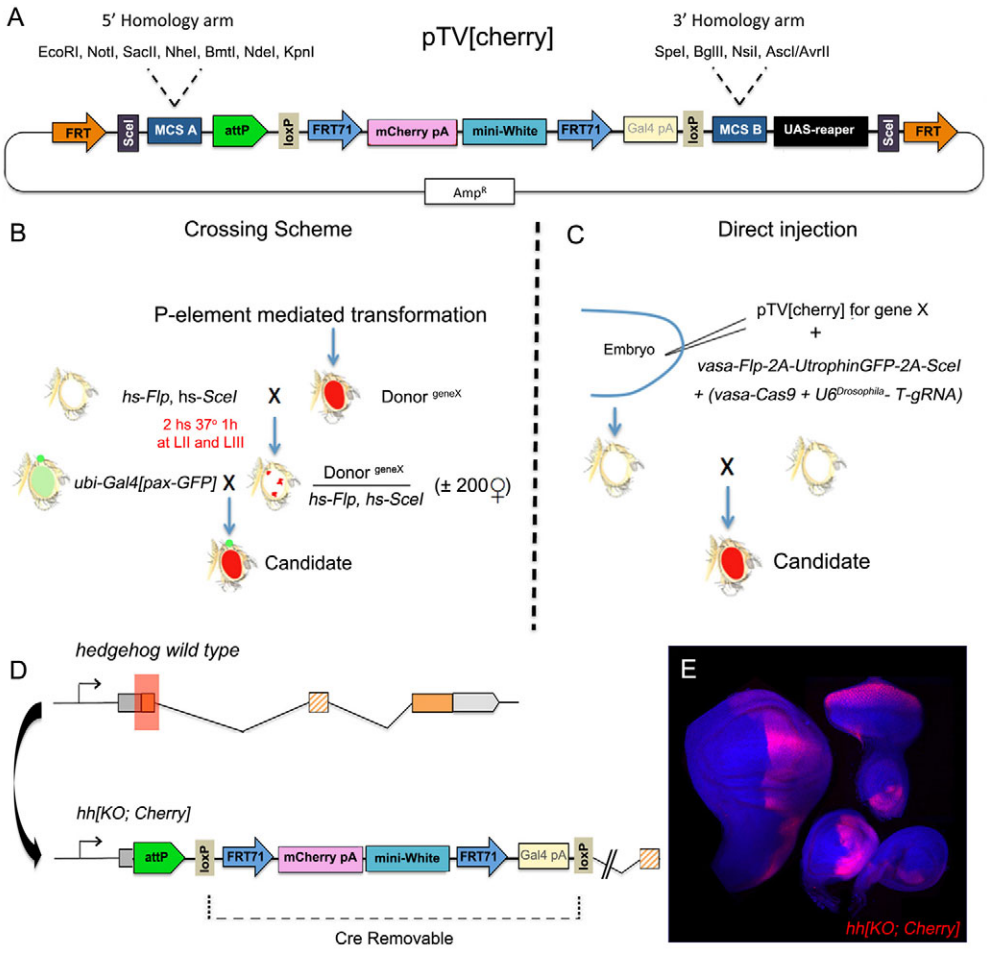

Fig. 1. A targeting vector and protocol for accelerated gene targeting. (A) The targeting vector, pTV^{Cherry}, showing its key features, including *cherry* and *mini-white*. The *cherry* cDNA and the *mini-white* were flanked by non-canonical FRTs (FRT71) and this was followed by Gal4-encoding sequences with the aim of enabling easy conversion to a Gal4 driver although this turned out not to be functional. (B) Gene targeting by crossing scheme. P-element-mediated transformation was used to generate the donor strain, which carries the targeting vector pTV^{Cherry}[geneX] at a random genomic location. The donor strain was crossed to flies carrying *hs-FLP* and *hs-I-SceI* to mobilise and linearize the donor construct (first cross). Typically, 200 adult female progeny with mottled eye colour (an indication that they carry the targeting vector and the *hs-FLP, hs-I-SceI* chromosome) were then crossed in pools of 15 to *ubiquitin-Gal4[3xP3-GFP]* (2d cross) to eliminate unwanted events and parental flies. The progeny of this second cross were then screened for red eyes and screening was terminated when a PCR-confirmed (on both sides) homologous recombinant was obtained. All the strains used were mutant at the endogenous *white* locus, allowing the *mini-white* of the targeting construct to be followed throughout. (C) Gene targeting by direct injection in embryos. The injected mix contained pTV^{Cherry}[geneX] and vasa-FUS (expressing *FLP* and *I-SceI*). For CRISPR-aided gene targeting, the mix included, in addition, U6-target-gRNA and vasa-Cas9. (D) Top: The *hedgehog* locus. Coding exons are orange, the first unaffected exon is shown as a hatched box, untranslated sequences are in grey and the region deleted by homologous recombination is marked by a red box. Bottom: The resulting allele, highlighting the Cre-excisable region that includes *cherry* and *mini-white*. (E) Cherry expression in imaginal discs from a *hedgehog[KO]* larva.

vector. For three out of the five constructs, accurate targeting was confirmed by PCR (Table 1; supplementary material Fig. S1). Therefore, pTV^{Cherry} is compatible with two strategies, one of which can be completed in 6 weeks (Fig. 1B,C).

As DSBs stimulate homologous recombination, we asked whether the frequency of targeting by injection could be increased by co-injecting a sequence-specific endonuclease. We chose CRISPR/Cas9 to induce DSBs because of the ease of construction of the necessary plasmids. Embryos were injected with a plasmid encoding Cas9 and another plasmid encoding a *wingless*-specific target gRNA, along with the pTV^{Cherry}[*wingless*] and the plasmid encoding *FLP* and *I-SceI* (Fig. 1C). In three independent experiments, confirmed homologous recombinants, which were easily recognised by the presence of *mini-white* and the pattern of Cherry expression, were obtained at a frequency of 1 per 500 injected embryos for *wingless* (Table 1). Although more genes will need to be targeted before an average frequency can be obtained, this result, along with the finding that CRISPR efficiently induces DSBs in *Drosophila* (Gratz et al., 2013), suggests that, in many cases, CRISPR will be an effective tool to increase the frequency of homologous recombination. Finally, we tested whether the length of the homology arms could be reduced. One homologous recombinant in *wingless* was retrieved following injection of a targeting vector containing 500 bp homologous arms and the CRISPR components. This result suggests that the long homology arms we used in previous experiments may not be necessary.

Post-targeting features and reintegration vectors

Besides the high frequency of homologous recombination, a key feature of pTV^{Cherry} is the transcriptional reporter activity provided by Cherry and hence the potential to quickly confirm true targeting events, as illustrated in Fig. 1E for *hedgehog*. Another important

Development (2013) doi:10.1242/dev.100933

Table 1. Frequency of gene targeting by crossing scheme and direct injection

	Homologous recombination by crosses		Homologous recombination by direct injection	
Gene	Confirmed KOs/ total flies screened	False positives	Confirmed KOs/ injected embryos	False positives
rhoGap102A	1/8000	1	0/3000	1
wingless	2/6000	0	2/3000	2
w/CRISPR	NR	NR	3/1500	0
w/CRISPR+500	NR	NR	1/500	0
hedgehog	2/4000	1	NA	NA
β-Catenin (*arm*)	2/4500	2	0/3000	3
dredd	2/6000	0	3/3000	2
sickle	2/2000	0	1/3000	1

Targeting by the crossing scheme was conducted as described in the legend of Fig. 1. Targeting by direct injection was achieved by injecting 3000 embryos (performed by Rainbow Transgenics). CRISPR-aided targeting was attempted for *wingless*, using the same homology arms and targeting vector as with the crossing scheme. This was carried out in three separate experiments involving 500 embryos each (one in house and two by Rainbow Transgenics). Each experiment led to one confirmed homologous recombinant. Finally, a targeting construct with 500 bp *wingless*-specific homology arms was injected along with CRISPR components (w/CRISPR+500) and one homologous recombinant was recovered following injection of 500 embryos. NA, not attempted; NR, not relevant.

benefit is the *attP* site. To expand the range of genetic manipulations that the *attP* site affords, we designed a range of reintegration vectors (Figs 2, 3; supplementary material Table S1). Using pGE-attB (Huang et al., 2009) as a starting point, we created RIV^{white} to allow reintegration of any cDNA (e.g. a wild-type form to confirm that the mutant phenotype associated with the KO is indeed due to destruction of the targeted gene). As RIV^{white} uses *mini-white* as a genetic marker, it requires Cre-mediated removal of *mini-white* and associated sequences from the targeted locus before reinsertion (supplementary material Fig. S3, Table S1). To avoid this requirement, we also constructed RIV^{Cherry}, which has the same polylinker but uses *3xP3-Cherry* as selectable marker (Fig. 2; supplementary material Fig. S3, Table S1). Thus, rescue by reintegration can be performed as soon as homologous recombination has been achieved (supplementary material Fig. S3). Using this procedure, reintegration of a *wingless* cDNA (including the 5′ and 3′ UTRs) in the *wingless* KO led to full rescue (not shown), suggesting that the *mini-white* and associated sequences remaining in the locus did not interfere with gene regulation. In the event that the cDNA did not rescue in this way, these exogenous sequences could easily be removed, along with the selectable marker of RIV, after reintegration (supplementary material Fig. S3). With

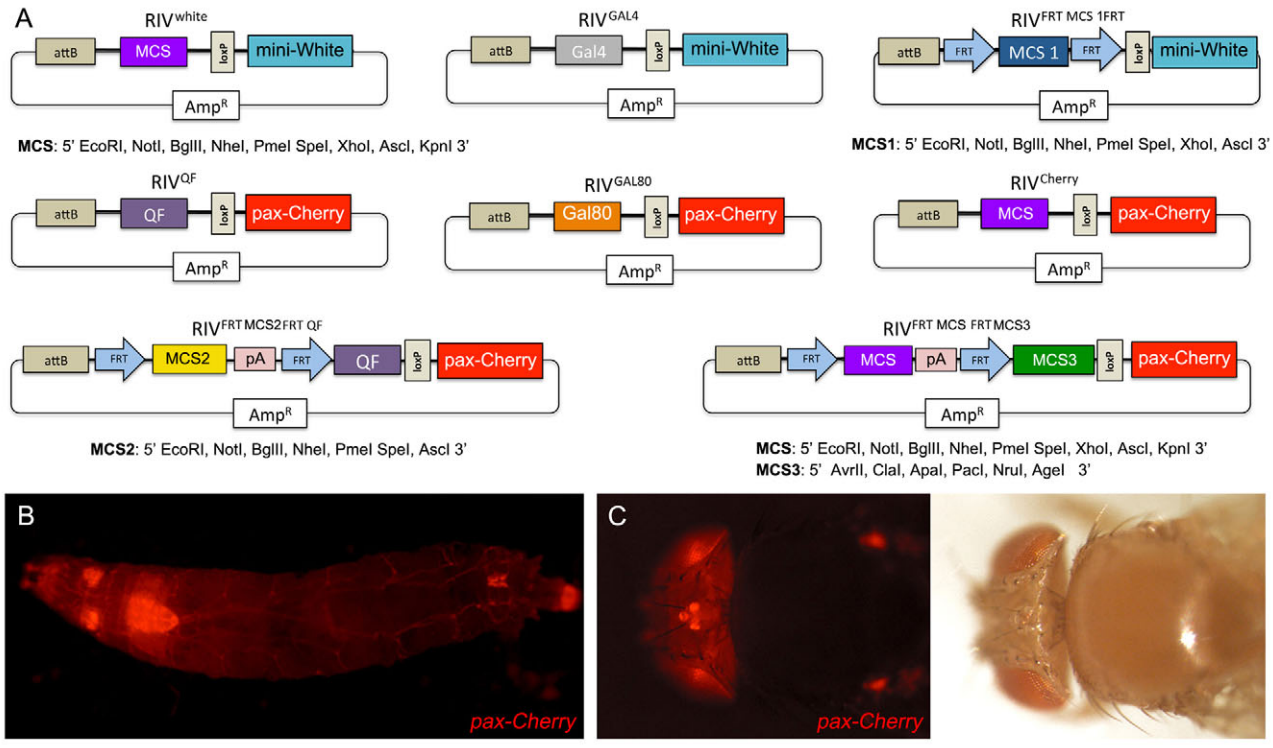

Fig. 2. A set of reintegration vectors (RIV): diagrams and features. (A) The key features of the reintegration vectors described in this paper. The restriction sites available for cloning (MCS) are listed below. Some of the vectors use *mini-white* and others use *3xP3-Cherry* (*pax-Cherry*) as genetic markers. (B,C) Low-magnification fluorescence micrographs show the suitability of *3xP3-Cherry* as a larval (B) and adult (C) markers. *RIV$^{FRT.Wg.FRT\ NRT\text{-}HA\text{-}Wg}$/Cyo[white+]* animals are shown.

Development (2013) doi:10.1242/dev.100933

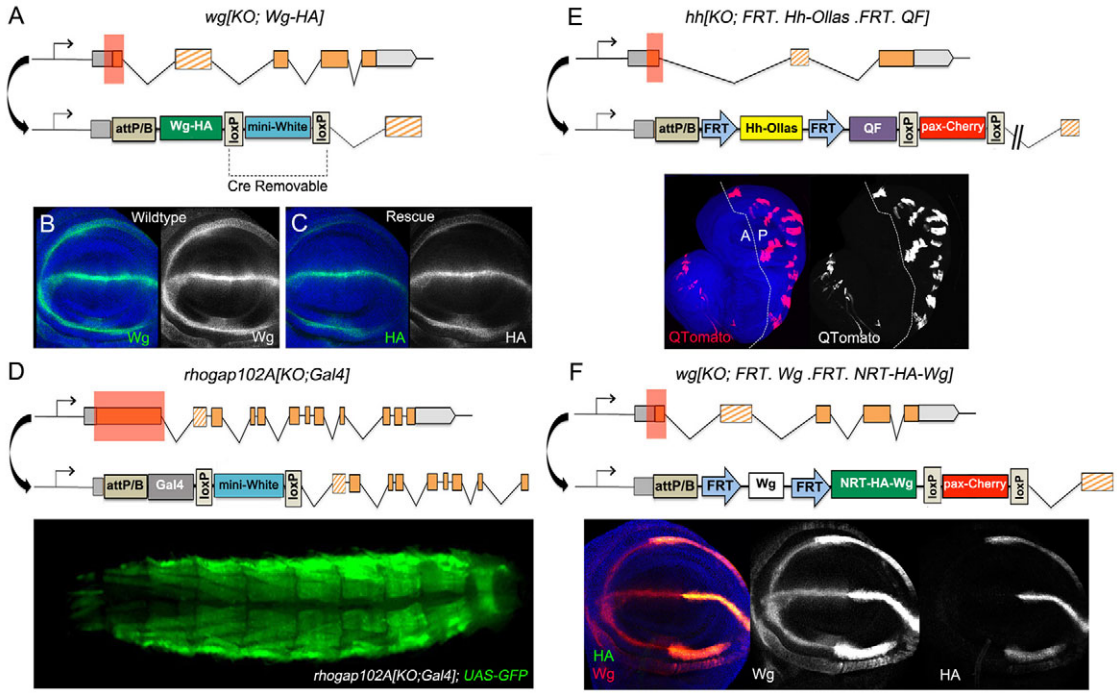

Fig. 3. Examples of reintegration vectors and their features. (A) The wild-type *wingless* locus before and after targeting by homologous recombination; Cre-mediated excision of the *cherry* and *mini-white* markers; and reintegration, via RIV^{white}, of a cDNA encoding HA-tagged Wingless. (B) Wild-type imaginal disc stained for anti-Wingless antibody. (C) Engineered allele expressing reintegrated HA-Wingless stained with anti-HA. (D) The *rhogap102A* locus before and after targeting by homologous recombination, Cre-mediated excision of the markers and reintegration of RIV^{Gal4}. Crossing such flies with UAS-GFP revealed that the *rhogap102A* locus is transcriptionally active in muscles. (E) The *hedgehog* locus before and after targeting by homologous recombination, Cre-mediated marker excision and reintegration of RIV$^{FRT.HhOllas.FRTQF}$. Larvae carrying the reintegrated allele *hs-FLP* and *QUAS-Tomato* were heat shocked to randomly excise the FRT cassette, thus triggering activation of *QUAS-Tomato* in a subset of the normal *hedgehog* expression domain. (F) Allele switching at the *wingless* locus. RIV$^{FRT.Wg.FRT\ NRT\text{-}HA\text{-}Wg}$ was reintegrated so that excision of the FRT cassette makes the locus stop expressing wild-type Wingless and start expressing HA-tagged NRT-Wingless instead. Excision was induced in the posterior compartment with *engrailed-Gal4*-driven *FLP*. Hence, HA immunostaining from NRT-HA-Wg is detected only in the posterior subset of *wingless*-expressing cells. DAPI stains nuclei in blue (all panels). Coding exons are orange, the first unaffected exon is shown as a hatched box, untranslated sequences are in grey and the region deleted by homologous recombination is marked by a red box.

RIV^{Cherry} or RIV^{white}, one can also re-integrate a tagged form of the cDNA, thus providing an assay for protein localisation without the need for an antibody, as illustrated for HA-tagged Wingless (Fig. 3A-C; supplementary material Table S1). If reintegration of the tagged cDNA rescues the mutant phenotype, as is the case for *wingless*, one can be confident that the tag does not significantly affect protein function and that its distribution is likely to reflect endogenous protein localisation. No transcription termination signals were included immediately downstream of the multiple cloning sites of RIV^{white} or RIV^{Cherry} (although the downstream markers, including *cherry* and *white*, evidently had one). As a rule, we favour use of a the endogenous 3′UTR of a gene for rescue experiments because it could contribute to RNA localisation (Lécuyer et al., 2007) and hence function. One other application of RIV^{white} or RIV^{Cherry} is that they can be used to integrate a reporter, e.g. encoding a fluorescent protein. To provide tools for misexpression specifically in the pattern of the targeted gene, we inserted a Gal4-encoding cDNA (Brand and Perrimon, 1993) in RIV^{white} and a QF-encoding cDNA (Potter et al., 2010) in RIV^{Cherry}, hence generating RIV^{Gal4} and RIV^{QF}, respectively. To illustrate the effectiveness of RIV^{Gal4}, it was reintegrated in the targeted *rhogap102A* gene. Crossing the resulting flies to *UAS-GFP* flies demonstrated that this gene is expressed in larval muscles (Fig. 3D). As an additional tool to control gene expression in defined patterns, we also created RIV^{Gal80}, which, upon reintegration in the targeted locus, allows gene-specific patterned repression of Gal4 (McGuire et al., 2003) (supplementary material Table S1; Fig. 2). Finally, we have created two reintegration vectors that flank the rescuing cDNA by two FRTs to render it excisable by *FLP* (Struhl and Basler, 1993) (RIV$^{FRT.MCS2.pA.FRT\ QF}$ and RIV$^{FRT.MCS.pA.FRT\ MCS3}$; supplementary material Table S1; Fig. 2). With these plasmids, gene activity can, in principle, be removed in an experimentally determined spatial and temporal pattern, even in post-mitotic or polyploid cells, a feat that has been nearly impossible so far. With RIV$^{FRT.MCS2.pA.FRT\ QF}$, following excision of the cDNA, the heterologous transcription activator QF becomes expressed and activates reporter genes such as QUAS-Tomato in these cells (Fig. 3E). In the related RIV$^{FRT.MCS.pA.FRT\ MCS3}$ vector, QF was replaced by a multiple cloning site distinct from MSC (MCS3) to allow insertion of another cDNA downstream of the one flanked by FRTs. This configuration can be used for cis-allele switching, whereby, for example, the wild-type cDNA can be replaced by a mutant allele (e.g. a point mutant) at specific times and places (Fig. 3F). Time-controlled tag switching (e.g. protein X with one tag in the first position and protein X with another tag in the second position) could also be potentially used to estimate protein turnover or trafficking between different subcellular compartments. Importantly, this can be achieved at physiological levels of expression under the control of the endogenous promoter.

DISCUSSION

We have described a vector that achieves high efficiency homologous recombination and provides immediate reporter gene

Development (2013) doi:10.1242/dev.100933

activity while retaining previously described features such as an easy-to-select genetic marker and an *attP* site. We have also improved on previous technology by devising a means of eliminating most illegitimate recombination events. Our targeting vector is compatible with three experimental strategies. One involves a cheap and effective two-generation crossing scheme following prior P-element-mediated transformation. The second involves embryonic injection of the targeting vector, along with an *I-SceI*-encoding plasmid, and selection in the F2 progeny. One might argue that *I-SceI* could be omitted from this protocol if the targeting vector were linearized prior to injection. However, in preliminary experiments, homologous recombinants were not obtained with linearized targeting vector (not shown). Moreover, previous experiments with small linear templates produced conflicting results (Beumer et al., 2008; Gratz et al., 2013). Because of these uncertainties, we have continued to inject supercoiled targeting vectors along with the *I-SceI*-encoding plasmid. Our third approach relies on the recent demonstration that DSBs stimulate homologous recombination in *Drosophila* (Beumer et al., 2008; Gratz et al., 2013). For this approach, we opted to induce DSBs with CRISPR and therefore co-injected four plasmids: the targeting vector, along with plasmids encoding a gene-specific guide RNA, Cas9 and *I-SceI*. We have chosen to express the guide RNA from a plasmid because this is more stable than RNA, an important consideration if the injection mix needs to be shipped to injection service providers. Importantly, our targeting vector is compatible with all the approaches described here so that if unexpected problems arose with CRISPR (e.g. off-target effects or recalcitrant genes) or if the generation of TALENs became much easier, our protocol could easily be adjusted without the need for new targeting constructs.

Despite the versatility of our targeting vector and protocols, there is undoubtedly room for further improvement. A seemingly higher frequency of CRISPR-aided homologous recombination was reported recently (Gratz et al., 2013). However, in this report, only a small double-stranded oligonucleotide was integrated, and recombinants therefore had to be selected by PCR from pools of candidate animals, a relatively laborious procedure. We suggest that the ease of screening with a visible eye marker, combined with the reporter activity of pTV^{Cherry} more than offsets the need to increase the number of embryos to inject. Nevertheless, it is likely that frequency improvements are possible. One possibility would be to create two DSBs instead of one (as reported by Gratz et al., 2013). This could be easily achieved by including a second gRNA-encoding plasmid in the injection mix. In addition, it is conceivable that the frequency of CRISPR-aided homologous recombination could be increased by using *vasa-Cas9* transgenic embryos as injection hosts or by adjusting the relative concentrations of the various injected plasmids. We have generated transgenic *vasa-Cas9* flies but they have yet to be tested. Targeting frequency might also be increased by introducing the guide RNA directly as RNA and not via a plasmid. However, as mentioned in results, we favour DNA injection because it can readily be outsourced. Another area of possible improvement concerns the size of homology arms as the relatively long arms (3-5 kb) that we have used in this study can be difficult to handle. As one experiment suggests (see Results), it is likely that 500 bp arms might suffice, although this will need further validation. We do not favour reducing the size further as this could lead to loss of specificity. Despite all the above suggestions for improvement, as they stand, our vector and protocols should help transform gene targeting by homologous recombination into a routine procedure in *Drosophila*, and perhaps in insect species that lack extensive genetic resources such as *Anopheles* or *Tribolium*.

Much of the genome editing possibilities described above are focused on the coding region. We have targeted the ATG-containing exon to preserve downstream regulatory elements (often found in the first intron) thus ensuring correct expression of reintegrated cDNAs. In all the cases we have studied so far, insertion of *Cherry* and *mini-white* did create a null allele. However, one must be aware that, following Cre-mediated excision of these elements, a truncated protein could become expressed from a downstream coding exon and restore partial function. Any such concern could be allayed by increasing the size of the deleted fragment appropriately. Another potential drawback of our approach is that the 'rescued locus' is not identical to that of the wild type. However, as long as full phenotypic rescue is achieved after reintegration, one can be confident that the subsequent analysis will be physiologically relevant. Although our targeting vector has been used so far to delete expressed exons, it is theoretically suitable for engineering any genomic sequence, including regulatory regions and enhancer elements. We envision that our targeting approach could be used routinely to study 1-2 kb regulatory element (a size that we have reliably deleted by homologous recombination). Large genomic regions (e.g. 20 kb) could be also engineered by two successive targeting operations to flank the region of interest with one *attP* site at each end. Mutated variants could then be reintegrated by recombination-mediated cassette exchange (RMCE) (Oberstein et al., 2005; Venken et al., 2011).

We have designed reintegration vectors that enable a variety of post-targeting applications. The basic reintegration vectors allow the insertion of a diverse set of reporters by co-injection with a plasmid encoding the PhiC31 integrase (Bateman et al., 2006; Bischof et al., 2007; Groth et al., 2004; Venken et al., 2006). They also provide a means of rapidly assessing the subcellular localisation and abundance of the protein product, as the reintegrated tagged cDNA should be expressed under endogenous control. Moreover, any targeted locus can be made to express Gal4, Gal80 and QF, thus widening the scope for sensitive reporter assays or misexpression studies. Finally, the combination of our FRT-containing reintegration vectors and the high efficiency of FLP-mediated cassette excision facilitate the removal of gene function in post-mitotic or polyploid cells at any time and in a tissue-specific manner. This, along with allele switching, opens up assays that have so far been nearly impossible.

MATERIALS AND METHODS

General *Drosophila* genetics

All the standard fly strains are described at http://flybase.org and all experiments were conducted at 25°C unless otherwise indicated. To induce clones in the domain of *hedgehog* expression, *y w hsp70-FLP UAS-GFP QUAST-tomato; RIV$^{FRT.HhOllas.FRT\ QF}$* larvae were cultured at 37°C for 50 minutes, 60±12 hours after egg laying. The resulting larvae were fixed and imaged 48 hours after the heat shock.

Immunostaining and microscopy

The following primary antibodies were used: rabbit anti-HA (1:1000, Cell Signaling Technology, C29F4) and mouse anti-Wingless (1:100, Hybridoma Bank). Secondary antibodies labelled with Alexa 488 or Alexa 555 (used at 1:200) were obtained from Molecular Probes. Imaginal discs were mounted in Vectashield with DAPI (Vector Laboratories). Fluorescence micrographs were acquired with a Leica SP5 confocal microscope. Embryo cuticles were prepared according to a standard protocol (Alexandre, 2008). Bright-field images from embryo cuticles were obtained with a Zeiss Axiophot2 microscope with an Axiocam HRC camera. Bright-field and confocal images were processed with Photoshop CS4 (Adobe).

Development (2013) doi:10.1242/dev.100933

Plasmid design and features

All PCRs were performed with Q5 High-Fidelity polymerase from New England Biolabs (NEB, M0492L). DNA synthesis was performed using Genewiz or Integrated DNA Technologies (IDT). The plasmids constructed for this study (most of which will be deposited at the DGRC, Indiana) are were: the targeting vector, pTV^{Cherry}; a plasmid expressing ubiquitous Gal4 to eliminate false positives – *ubiquitin-Gal4[3xP3-GFP]*; various reintegration vectors – RIV^{white}, RIV^{Gal4}, RIV$^{FRT.MCS2.pA.FRT\ QF}$, RIV^{Cherry}; RIV^{QF}, RIV^{Gal80}, RIV$^{FRT.MCS.pA.FRT\ MCS3}$, RIV$^{FRT.MCS1.FRT}$, RIV$^{FRT.HhOllas.FRT\ QF}$, RIV$^{WgHA/white}$ and RIV$^{FRT.Wg.FRT\ NRT\text{-}HA\text{-}Wg}$; a plasmid expressing Flp, Utrophin-GFP and *I-SceI* from a single transcript using the 2A peptide system (Trichas et al., 2008) (note that only *I-SceI* is relevant to the present work; see Results) – vasa-FUS; a plasmid expressing a *wingless*-specific gRNA – U6-*wingless*-gRNA; a plasmid expressing a *sickle*-specific gRNA – U6-sickle-gRNA; a plasmid for expression of any gene-specific gRNA following insertion of a suitable double-stranded oligonucleotide – U6-*Bsa*I-gRNA; and a plasmid expressing Cas9 under the control of the *vasa* germline-specific – vasa-Cas9. Details on their construction are provided in supplementary material Appendix S1.

Gene targeting

For targeting with the crossing scheme, the targeting vector, which was modified to contain the appropriate homology arms (here 3-5 kb arms were used but smaller arms are likely to suffice), was introduced at a random genomic locations by P-element-mediated transformation (Rainbow Transgenics, Bestgene or in house). Transformants (not necessarily mapped or homozygosed) were crossed to *hs-FLP, hs-I-SceI* flies (Bloomington stock number 25679 or 25680) and the resulting larvae were heat-shocked at 48 and 72 hours after egg laying (AEL) for 1 hour at 37°C. Approximately 200 adult females with mottled red eyes (indicating the presence of the targeting vector and the transgene carrying *hs-FLP* and *hs-I-SceI*) were crossed in pools of 15 (this number was chosen to maximise the number of progeny for the vial size we use) to *ubiquitin-Gal4[3xP3-GFP]* males and the progeny was screened for the presence of red-eyed flies. The *ubiquitin-Gal4[3xP3-GFP]* transgene was subsequently removed by selecting against the presence of GFP in the ocelli. In most, but not all, cases screening was stopped after a homologous recombinant was confirmed. Gene targeting by direct injection was performed as for P-element-mediated transformation, except that the injection mix contained the targeting vector (700 ng/µl) and vasa-FUS (300 ng/µl). DNA was prepared with the PureLink Midi prep Kit (Invitrogen, K2100-16). For each targeting experiment, 3000 embryos were injected by Rainbow Transgenics. The resulting flies were crossed to *white^{1118}* flies and red-eyed flies were selected from the progeny. For CRISPR-aided homologous recombination, the injection mix contained the targeting vector (600 ng/µl), vasa-FUS (200 ng/µl), U6-target-gRNA (150 ng/µl) and vasa-Cas9 (150 ng/µl), all as supercoiled plasmids. We have obtained evidence that, using this method, 500 bp homology arms might suffice.

Detailed protocol for accelerated homologous recombination in *Drosophila*

Materials

- Plasmids (DGRC, Indiana): pTV^{Cherry}, vasa-FUS, U6-*Bsa*I-gRNA and vasa-Cas9.
- *Drosophila* genomic DNA or bacterial artificial chromosome (BAC) containing the genomic region to be targeted (http://www.pacmanfly.org).
- *Drosophila*: a healthy *white* mutant strain.

Methods

- Determine genomic region to be deleted, identify Cas9 target sequence ($GN_{20}GG$) within it, obtain oligonucleotides, anneal and ligate into *Bsa*I site of U6-*Bsa*I-gRNA to generate U6-geneX-gRNA. See examples below with Cas9 target sequence in bold and sequence to insert in U6-*Bsa*I-gRNA underlined. The transcription start site is in italics.

Example 1 (CRISPR site on the top strand)

Genomic sequence

5′ ...GCAAACAGAAATCCG**GAAGGGGCCGGGGCTCCATGTGG**TG... 3′

3′ ...CGTTTGTCTTTAGGCCTTCCCCGGCCCCGAGGTACACCAC... 5′

Pair of oligonucleotides to order (1st G from Cas9 site +19 bp)

5′ TTC**GAAGGGGCCGGGGCTCCATG** 3′

3′ TTCCCCGGCCCCGAGGTACCAAA 5′

Sequence after ligation in U6-*Bsa*I-gRNA

...*TATAT*AGGTATGTTTTCCTCAATACTTC**GAAGGGGCCGGGGCTCCATG**–GTTTTAGAGCTAGAAATAG...

Example 2 (CRISPR site on the bottom strand)

Genomic sequence

5′ ...GCTCCCGTTTGTCTTTAGGCCTTCCCCGGCCCCGAGGTAC... 3′

3′ ...CGA**GGGCAAACAGAAATCCGGAAGGG**GCCGGGGCTCCATG... 5′

Pair of oligonucleotides to order (1st G from Cas9 site +19 bp)

5′ TTC**GGGAAGGCCTAAAGACAAAC** 3′

3′ CCTTCCGGATTTCTGTTTGCAAA 5′

Sequence after ligation in U6-*Bsa*I-gRNA

...*TATAT*AGGTATGTTTTCCTCAATACTTC**GGGAAGGCCTAAAGACAAAC**–GTTTTAGAGCTAGAAATAG...

- Amplify homology arms by PCR (>500 bp) using a suitable BAC or genomic DNA as a template and ligate into multiple cloning sites of pTV^{Cherry} to generate pTV^{Cherry}[GeneX].
- Prepare injection mix comprising (all plasmids), pTV^{Cherry}[GeneX] (600 ng/µl), U6-geneX-gRNA (150 ng/µl), vasa-FUS (200 ng/µl) and vasa-Cas9 (150 ng/µl), and inject (can be subcontracted) into *white* mutant embryos.
- If funds permit, generate in parallel pTV^{Cherry}[GeneX] transformants by P-element-mediated transformation. These flies could be used in a backup CRISPR-independent protocol, as described in Fig. 1B.
- Cross all the adults arising from injected embryos individually or in small groups to *white* mutants and screen for *white*$^{+}$ flies in the progeny (can also be subcontracted).

Acknowledgements

Sebastian Gerety (Sanger Institute) provided CRISPR/Cas9-related advice and tools. Konrad Basler (Zurich), Chris Potter (Johns Hopkins), Yang Hong (Pittsburgh) and Sean Megason (Harvard) provided plasmids.

Competing interests

The authors declare no competing financial interests.

Author contributions

L.A.B.-L., C.A. and J.-P.V. contributed equally to the conception of the work and the preparation of the manuscript. Most experimental work was performed by L.A.B.-L. and C.A. A.M. and L.P. contributed important experimental results.

Funding

This work was supported by the UK Medical Research Council, by an European Research Council grant [WNTEXPORT] from the European Union to J.P.V. and by a Sir Henry Wellcome post-doctoral fellowship to L.A.B. [082694/Z/07/Z]. Deposited in PMC for immediate release.

Supplementary material

Supplementary material available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.100933/-/DC1

References

Alexandre, C. (2008). Cuticle preparation of Drosophila embryos and larvae. *Methods Mol. Biol.* **420**, 197-205.

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

Bateman, J. R., Lee, A. M. and Wu, C. T. (2006). Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. *Genetics* **173**, 769-777.

Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug, R. G., II, Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y. H. et al. (2012). In vivo genome editing using a high-efficiency TALEN system. *Nature* **491**, 114-118.

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

Bischof, J., Maeda, R. K., Hediger, M., Karch, F. and Basler, K. (2007). An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. USA* **104**, 3312-3317.
Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415.
Capecchi, M. R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat. Rev. Genet.* **6**, 507-512.
Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. A., Somia, N. V., Bogdanove, A. J. and 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.
Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.-W. and Xi, J. J. (2013). Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* **23**, 465-472.
Cho, S. W., Kim, S., Kim, J. M. and Kim, J.-S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230-232.
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. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823.
DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J. and Church, G. M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336-4343.
Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K. and Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822-826.
Gaj, T., Gersbach, C. A. and Barbas, C. F., III (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397-405.
Gong, W. J. and Golic, K. G. (2003). Ends-out, or replacement, gene targeting in Drosophila. *Proc. Natl. Acad. Sci. USA* **100**, 2556-2561.
Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., Wildonger, J. and O'Connor-Giles, K. M. (2013). Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**, 1029-1035.
Groth, A. C., Fish, M., Nusse, R. and Calos, M. P. (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics* **166**, 1775-1782.
Horn, C., Jaunich, B. and Wimmer, E. A. (2000). Highly sensitive, fluorescent transformation marker for Drosophila transgenesis. *Dev. Genes Evol.* **210**, 623-629.
Huang, J., Zhou, W., Watson, A. M., Jan, Y.-N. and Hong, Y. (2008). Efficient ends-out gene targeting in Drosophila. *Genetics* **180**, 703-707.
Huang, J., Zhou, W., Dong, W., Watson, A. M. and Hong, Y. (2009). From the Cover: Directed, efficient, and versatile modifications of the Drosophila genome by genomic engineering. *Proc. Natl. Acad. Sci. USA* **106**, 8284-8289.
Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J. and Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227-229.
Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T. R., Tomancak, P. and Krause, H. M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**, 174-187.
Liu, J., Li, C., Yu, Z., Huang, P., Wu, H., Wei, C., Zhu, N., Shen, Y., Chen, Y., Zhang, B., et al. (2012). Efficient and specific modifications of the drosophila genome by means of an easy TALEN strategy. *Acta Genetica Sinica* **39**, 209-215.
Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. and Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826.
McGuire, S. E., Le, P. T., Osborn, A. J., Matsumoto, K. and Davis, R. L. (2003). Spatiotemporal rescue of memory dysfunction in Drosophila. *Science* **302**, 1765-1768.
Oberstein, A., Pare, A., Kaplan, L. and Small, S. (2005). Site-specific transgenesis by Cre-mediated recombination in Drosophila. *Nat. Methods* **2**, 583-585.
Potter, C. J., Tasic, B., Russler, E. V., Liang, L. and Luo, L. (2010). The Q system: a repressible binary system for transgene expression, lineage tracing, and mosaic analysis. *Cell* **141**, 536-548.
Rong, Y. S. and Golic, K. G. (2000). Gene targeting by homologous recombination in Drosophila. *Science* **288**, 2013-2018.
Struhl, G. and Basler, K. (1993). Organizing activity of wingless protein in Drosophila. *Cell* **72**, 527-540.
Trichas, G., Begbie, J. and Srinivas, S. (2008). Use of the viral 2A peptide for bicistronic expression in transgenic mice. *BMC Biol.* **6**, 40.
Venken, K. J. T., He, Y., Hoskins, R. A. and Bellen, H. J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. *Science* **314**, 1747-1751.
Venken, K. J. T., 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. (2011). MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. *Nat. Methods* **8**, 737-743.
Weng, R., Chen, Y.-W., Bushati, N., Cliffe, A. and Cohen, S. M. (2009). Recombinase-mediated cassette exchange provides a versatile platform for gene targeting: knockout of miR-31b. *Genetics* **183**, 399-402.
Xie, H. B. and Golic, K. G. (2004). Gene deletions by ends-in targeting in Drosophila melanogaster. *Genetics* **168**, 1477-1489.
Zhou, W., Huang, J., Watson, A. M. and Hong, Y. (2012). W:Neo: a novel dual-selection marker for high efficiency gene targeting in Drosophila. *PLoS ONE* **7**, e31997.