

Efficient chromosomal gene modification with CRISPR/*cas9* and PCR-based homologous recombination donors in cultured *Drosophila* cells

Romy Böttcher^{1,†}, Manuel Hollmann^{1,†}, Karin Merk^{1,†}, Volker Nitschko^{1,†},
Christina Obermaier^{1,†}, Julia Philippou-Massier^{1,†}, Isabella Wieland^{1,†}, Ulrike Gaul^{1,2} and
Klaus Förstemann^{1,*}

¹Gene Center and Dept. of Biochemistry, Ludwig-Maximilians-Universität, Feodor-Lynen-Str. 25, D-81377 München, Germany and ²Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-Universität, Feodor-Lynen-Str. 25, D-81377 München, Germany

Received February 24, 2014; Revised March 26, 2014; Accepted March 26, 2014

ABSTRACT

The ability to edit the genome is essential for many state-of-the-art experimental paradigms. Since DNA breaks stimulate repair, they can be exploited to target site-specific integration. The clustered, regularly interspaced, short palindromic repeats (CRISPR)/*cas9* system from *Streptococcus pyogenes* has been harnessed into an efficient and programmable nuclease for eukaryotic cells. We thus combined DNA cleavage by *cas9*, the generation of homologous recombination donors by polymerase chain reaction (PCR) and transient depletion of the non-homologous end joining factor *lig4*. Using cultured *Drosophila melanogaster* S2-cells and the phosphoglycerate kinase gene as a model, we reached targeted integration frequencies of up to 50% in drug-selected cell populations. Homology arms as short as 29 nt appended to the PCR primer resulted in detectable integration, slightly longer extensions are beneficial. We confirmed established rules for *S. pyogenes cas9* sgRNA design and demonstrate that the complementarity region allows length variation and 5'-extensions. This enables generation of U6-promoter fusion templates by overlap-extension PCR with a standardized protocol. We present a series of PCR template vectors for C-terminal protein tagging and clonal *Drosophila* S2 cell lines with stable expression of a myc-tagged *cas9* protein. The system can be used for epitope tagging or reporter gene knock-ins in an experimental setup that can in principle be fully automated.

INTRODUCTION

It has taken only a surprisingly short time between the initial discovery of CRISPR/*cas* systems as a phage defense mechanism (1) based on RNA-programmed (2), sequence-specific nucleases (3) and the development of a specific subtype into a versatile tool for genome editing (4–6). It has since been successfully applied to budding yeast (7), *Caenorhabditis elegans* and other nematodes (8–15), *Drosophila* (16–21), *Arabidopsis*, *Nicotiana*, *Sorghum* and *Oryza* (22–26), zebrafish (27–33), *Xenopus* (34), mice (32,35–41), rats (42), rabbits (43), beef (44), cynomolgus monkey (45) and human cells (46–50)—all within little more than a year. There are few examples of comparable success of a new technology. The broad applicability—from yeast to plants and mammals—indicates that the system is indeed highly efficient despite the fact that it has not evolved to cleave DNA associated with eukaryotic nucleosomes. Researchers are now working to improve cleavage specificity of the system (51–55) and to make use of the programmable, sequence-specific binding for other purposes than DNA cleavage (56–62).

Several of the above-cited studies already made use of CRISPR/*cas9* induced DNA breaks to stimulate repair via an experimentally provided homologous recombination (HR) donor construct. These editing tools were either large plasmid-based constructs with long regions of sequence homology (kb range) or single-stranded oligonucleotides with a short region of sequence homology. The plasmid constructs usually template repair efficiently and can transfer large tag sequences, but their generation is time-consuming. Single-stranded oligonucleotides are conveniently produced via chemical synthesis but due to size restraints their 'coding capacity' is limited; they are most useful for specifying point mutations in active sites or introduction of e.g. *loxP*

*To whom correspondence should be addressed. Tel: +49 89 218076912; Fax: +49 89 218076945; Email: Foerstemann@genzentrum.lmu.de

†The authors wish it to be known that, in their opinion, the first seven authors should be regarded as Joint First Authors.

or *attP* sites (reviewed in (63)). An approach that unites the convenience of oligonucleotide ordering with the capacity to introduce large tags has been developed in budding yeast (64): a series of plasmid templates combines e.g. epitope tags and a selectable marker, which are then amplified with flanking primers that specify the desired integration site via appended homology arms. The presumably lower integration efficiency of the short homology arms is outweighed by the very efficient selection that is possible with microorganisms.

Inspired by the success of the CRISPR/*cas9* technology and the ease of protein tagging available in budding yeast, we attempted to use polymerase chain reaction (PCR)-based homology donors for genome editing in cultured *Drosophila* S2 cells. Efficient selection of stably transformed cells via a blasticidin resistance gene and stable expression of the *cas9* nuclease allowed us to readily recover cells with targeted integration of a donor construct at the phosphoglycerate kinase (PGK) locus flanked by as little as 29 nt of sequence homology on either end. Repair via the non-homologous end joining (NHEJ) pathway is a competing reaction for this purpose, eventually leading to mutation of the CRISPR target site without integration of the donor construct. Consistent with this notion, we find that depletion of the *lig4* gene (65) via RNA interference prior to induction of the DNA break substantially increased the proportion of cells harbouring the desired integrates, reaching a level of up to 50% of the blasticidin resistant cell population. Similar effects have been observed in *lig4* mutant *Drosophila* embryos with Zn-finger induced DNA double-strand breaks (DSBs) (66). For many purposes this level may already be sufficient; furthermore, we found it straightforward to derive cell clones with uniform expression by serial dilution. Compared with transient transfections or stable cell lines generated by conventional approaches, our technique is not only more convenient but also avoids potential artifacts due to heterogeneous and exaggerated expression levels of the tagged protein. The approach we are presenting here is based on transfection of PCR products and *in vitro* transcribed guide RNA. The latter can be replaced by U6-promoter-sRNA template fusions also generated by PCR according to a standardized protocol. However, in this case spontaneous integration of the gRNA expression cassette may lead to permanent activation of the *cas9* nuclease and, consequently, a high potential for off-target cleavage events.

MATERIALS AND METHODS

Molecular biology techniques

The cleavable double-GFP reporter was generated by inserting a *NotI/XbaI* digested PCR product of green fluorescent protein (GFP) obtained with primers 5'-atGCGGCCGCGTGAGCAAGGGCGAGGAGCT-3' and 5'-TAtctagaggccgtTACTTGTACAGCTCGTCCATGC-3' into *NotI/XbaI* cut pKF63 (67). The resulting plasmid was opened at the *NotI*-site in between the two GFP moieties and the annealed oligonucleotides 5'-ggccTAGGGATAACAGGGTAATgc-3' and 5'-ggccgcATTACCCTGTTATCCCTA-3' were inserted; the orientation was determined by sequencing. The final

plasmid, pKF251, has a ubiquitin promoter, a fully functional copy of myc-GFP followed by the *I-SceI* recognition site and a second GFP that lacks an N-terminal methionine. We amplified the *I-SceI* ORF from a fly strain harbouring a heat-shock promoter driven *I-SceI* transgene (68) with primers 5'-caGGTACCTAATCCAAAATGTTTCATGCC TTCTTCTTTTTCC-3' and 5'-atGCGGCCGCTTATTT CAGGAAAGTTTCGGAGGAG-3', digested the PCR product with *KpnI/NotI* and inserted the fragment into *KpnI/NotI*-cut pCASPER tub (69) yielding pKF257. This vector leads to expression of the *I-SceI* nuclease under the control of the α -tub84 promoter.

The *cas9* gene with humanized codon bias (47) was amplified with primers 5'-ATggtaccTAATCCAAAatgga acagaaactgattagcgaagaagacctgGACAAGAAGTACTC CATTGG-3' (appending a myc-tag at the N-terminus) and 5'-atgcgccgcTCACACCTTCTTCTTCTTG-3', digested with *KpnI/NotI* and ligated into correspondingly cleaved pKF257 (see above). Furthermore, an *attB* sequence was inserted into a unique *NdeI* site in this plasmid with annealed oligos 5'-TATGgggtgccaggcgtgccttgggctccccggcgcgta-3' and 5'-TAtacgcgccggggagcccaaggcagccctggcaccCA-3' resulting in plasmid pRB14 (expression of myc-*cas9*).

To generate a vector that allows for expression of the site-specific recombinase FLP from *Saccharomyces cerevisiae*, the gene was PCR amplified from yeast genomic DNA (strain W303) using primers 5'-atggtaccATGCCACAATTTGGTATATTATGTAAAAC-3' and 5'-atgcgccgcTTATATGCGTCTATTTATGTAG GATGAAAG-3', then inserted into *BamHI/NotI* cut pKF63 yielding plasmid pMH5. The insert was sequence verified.

To generate sgRNA templates for *in vitro* transcription via PCR, the oligonucleotide 5'-GTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGC-3' served as PCR template with a sense primer containing the T7 promoter and targeting sequence (e.g. 5'-taatacgaactataGCCTAGGGATAACAGGGTAATGg ttttagact-3') and 5'-GCACCGACTCGGTGCCACT-3' as antisense primer. For the 3'-extended CRISPR's, oligo 5'-GTGAGCAAGGGCGAGGAGgcaccgactcgtgcccact-3' served as antisense PCR primer. Details on CRISPR target site oligonucleotides employed are given in Figure 2A and the supplementary information. *In vitro* transcription was performed as previously described for dsRNA (70), the sgRNA products were purified via a Qiagen PCR purification kit. PCR amplification of the U6-C promoter with a T7-extension was achieved with oligonucleotides 5'-GCTCACCTGTGATTGCTCCTAC-3' and 5'-atagtag tctgattaAACGACGTTAAATTGAAAATAGGTCTA-3'. The PCR product was cloned into pJet1.2 resulting in plasmid pRB17 and sequence verified. Overlap-extension PCR was performed with 1 μ l of sgRNA *in vitro* transcription template (see above) and 1 μ l of a 10 ng/ μ l dilution of pRB17 per 50 μ l PCR. The primers for this PCR were 5'-GCTCACCTGTGATTGCTCCTAC-3' and 5'-gcttattctcAAAAAAGCACCGACTCGGTGCCACT-3' (to introduce a pol-III termination signal at the end).

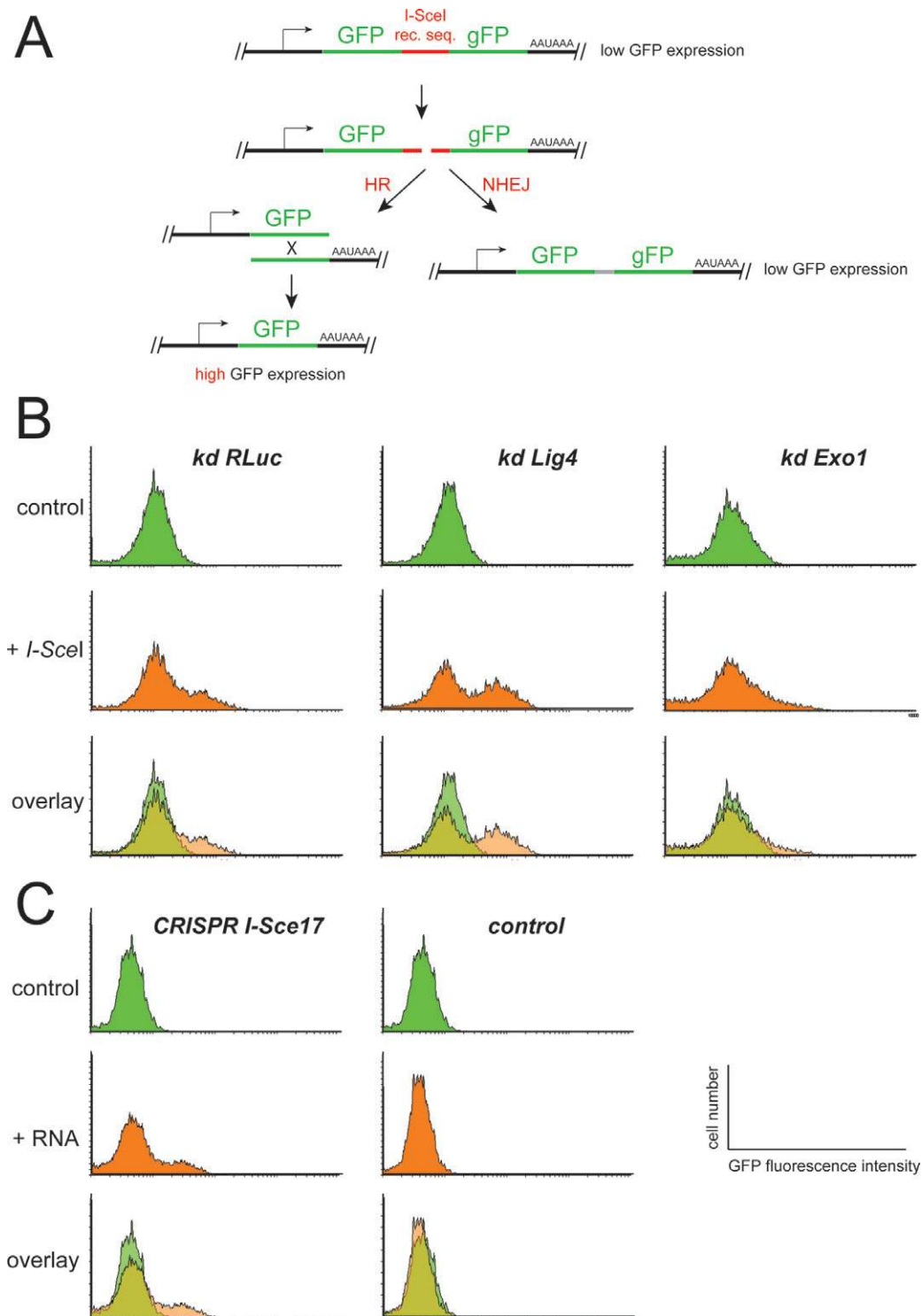


Figure 1. *I-SceI* and CRISPR-mediated cleavage of chromosomal DNA occur with comparable efficiency. (A) Schematic representation of the double-GFP reporter construct. Note that the first copy of GFP is complete and therefore functional, and the second copy only lacks the initiating methionine (symbolized by the lower case g). The non-rearranged reporter leads to low GFP fluorescence in clonally selected, stable cell lines. Upon cleavage at the intervening *I-SceI* recognition sequence, repair may occur either via homologous recombination (HR, left side) or non-homologous end joining (NHEJ, right side). While HR rearranges the locus and results in high GFP expression, NHEJ leads to short deletions but retains the low level of GFP fluorescence. (B) Depletion of *lig4* (an essential NHEJ factor) and *exo1* (an essential HR factor) followed by transfection of an *I-SceI* nuclease expression vector demonstrates the preferred outcome of reporter fluorescence when either pathway is suppressed. Knock-down of *Renilla* luciferase served as a control. (C) In clonally selected stable cell lines that express a myc-tagged version of the *cas9* nuclease protein in addition to the reporter construct, transfection of an *in vitro* transcribed, cognate CRISPR RNA (see Figure 2A for sequence details, *I-Sce17* refers to the 17-nt-long CRISPR targeting region) leads to a comparable extent of DNA cleavage as the *I-SceI* nuclease itself (as judged by the proportion of GFP^{high} cells). A control RNA (CRISPR target sequence 5'-GCGGTGGACCAGCTGCAGC-3') that does not target *cas9* to our reporter did not lead to detectable cleavage and rearrangement.

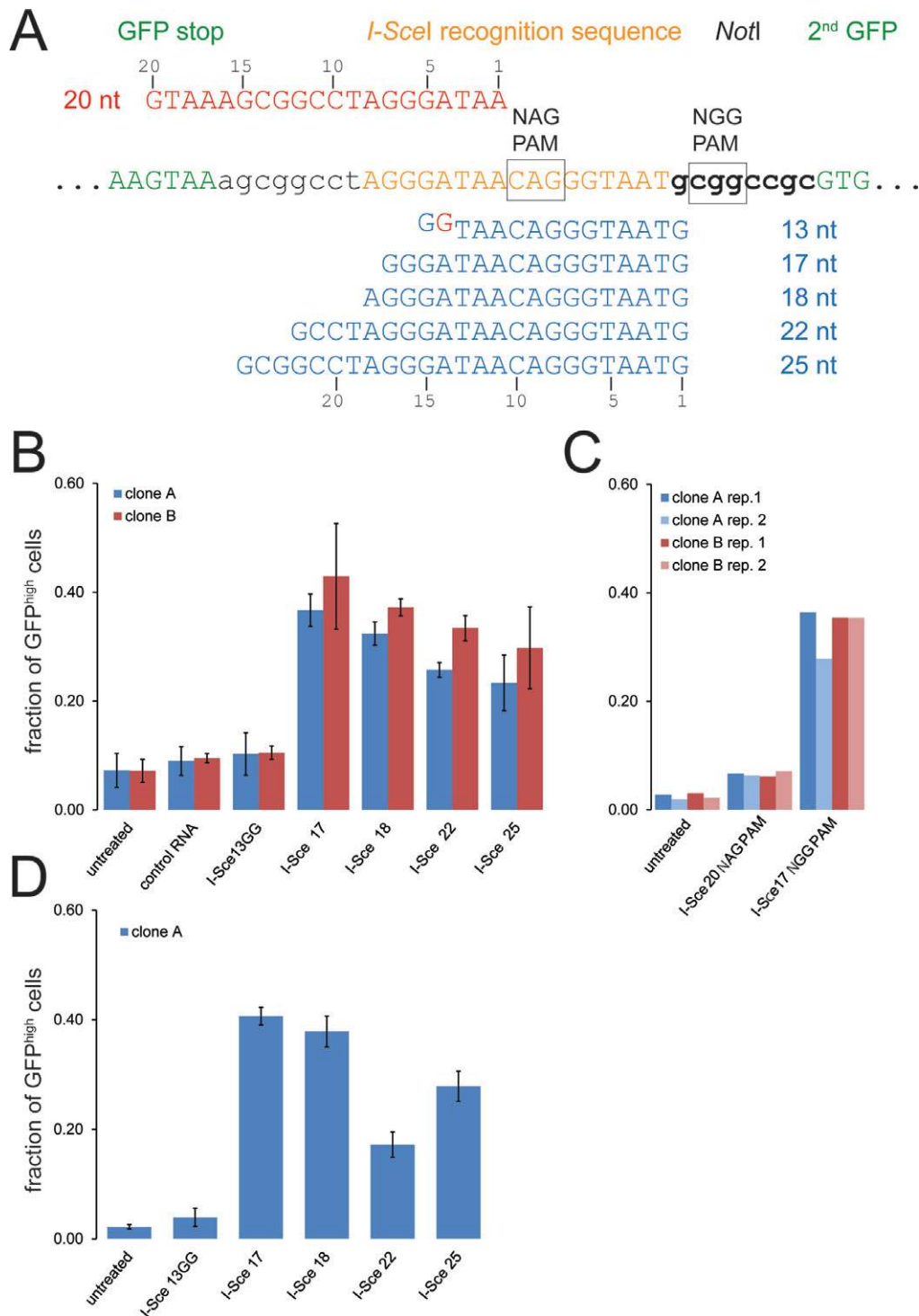


Figure 2. Quantitative assessment of *cas9*-mediated cleavage in living cells using various CRISPR RNA guides. (A) Sequence detail of the *I-SceI* recognition site region in our reporter construct. The length variant CRISPR RNAs with an NGG trinucleotide protospacer-associated motif (PAM) are shown in blue below the reporter sequence and the CRISPR RNA for testing cleavage efficiency at the 5'-NAG-3' trinucleotide PAM is shown in red above the reporter sequence. (B) Quantitative analysis of CRISPR cleavage activity according to targeting sequence length using our HR reporter system. The proportion of GFP^{high} cells was determined via two-dimensional analysis of fluorescence-activated cell sorting data (side scatter vs. GFP fluorescence) as this enables more reliable separation of the two populations. Two independent cell clones expressing *myc-cas9* and the double-GFP reporter were analyzed; the data are presented as the mean \pm SD of three independent biological replicates. (C) Quantitative analysis of CRISPR cleavage activity with an NAG trinucleotide PAM. The experiment was performed essentially as described in B, the *I-SceI*7 NGG CRISPR serves as comparison. Cleavage activity towards an NAG PAM is detectable but occurs with clearly lower efficiency. Two cell clones were tested in two independent biological replicates each. (D) Extending the CRISPR RNA at its 3'-end does not impair cleavage activity. The experiment was performed and analyzed as in B but a CRISPR RNA with a 3'-extension harboring 18 nt of sequence homology downstream of the CRISPR target site was employed. This extension did not impair cleavage activity (compare with B), but also did not rescue the apparent defect of the *I-SceI*3GG RNA construct either.

A *copia-Bsd-eGFP* fusion gene was constructed by amplifying the *copia* promoter from pDrBB2 (Addgene) with primers 5'-GCgaattcCTGTTGGAATATACTATTCAAC-3' and 5'-GCggtaccTCCTTTCTTAATAAATAAATAAATAG-3', the *bsd* gene from pCMV/Bsd (Invitrogen) with primers 5'-GCaccggtATGGCCAAGCCTTTGTCTCA-3 and 5'-GCccatggtGCCCTCCACACATAACCA-3' and the eGFP reporter from pEGFP-N1 (Addgene) with primers 5'-GCccATGGTGAGCAAGGGCGAG-3' and 5'-GCgctagcTTACTTGTACAGCTCGTCCA-3', then inserting them via restriction digest into a pUC19 derivative.

Template plasmids for HR-mediated protein tagging were generated by amplifying the *copia-Bsd* portion of the fusion gene described above with primers 5'-GATTATAAAGATGATGATGATAAAactagtGAAGTTCCTATACTTTCTAGAGAATAGGAACCTTCAAATAAAacatatgCTGTTGGAATATACTATTCAAC-3' and 5'-atCATATGTTAGAAACAAATTTATTTTTAAAGTTTTATTTTTAATAATTTctaGCCCTCCACACATAACCAGA-3' to remove the GFP ORF, add a 5'-FRT sequence, a stop codon and a short 3'-UTR. The product was re-amplified with the same sense primer and 5'-GAAGTTCCTATTCTCTAGAAAGTATAGGAACCTTccatagttagaacaatttttaaaag-3' to append a downstream FRT site, then inserted via blunt-end cloning into pJet2.1. The orientation was determined by sequencing. This plasmid served as the basis for tag insertion between an *Xho*I and *Spe*I site. We used primers 5'-tactegagATGGTGAGCAAGGGCGAGGAG-3' and 5'-taactagtTTACTTGTACAGCTCGTCCATG-3' to amplify the coding sequences of EGFP resulting in plasmid pMH3. A double FLAG-tag template was constructed by ligating the annealed oligonucleotides 5'-tcgagGATTATAAAGATGATGATGATAAAatccggagccGATTATAAAGATGATGATGATAAATGAa-3' and 5'-ctagtTCATTTATCATCATCATCTTTATAATCggtccggaTTTATCATCATCATCTTTATAATCc-3' into the base vector resulting in plasmid pMH4. The TwinStrep-tag sequence was inserted via the annealed oligonucleotides 5'-tcgagAGCGCTTGGAGCCACCCGCAGTTCGAGAAAGGTGGAGGTTCCGGAGGTGGATCGGGAGGTGGA TCGTGGAGCCACCCGCAGTTCGAAAAATGAa-3' and 5'-ctagtTCATTTTTCGAACTGCGGGTGGCTCCACGATCCACCTCCCGATCCACCTCCGGAACCTCCACCTTTCTCGAACTGCGGGTGGCTCCAAGCGCTc-3' resulting in plasmid pIW1.

PCR amplification of targeting constructs (HR donors) was performed with one of the plasmids described above as template and the primers annealing with their 3'-end sequences of ...ggatcttccggatggctcgag-3' and ...gaagttcctattctctagaaagttaggaacttccatag-3'. Further details are provided in Figure 3 and the supplementary information. For PGK targeting, the primers were 5'-CGTCTCCACCGAGGCGGCGCTTCGCTGGA GCTCCTGGAGGGCAAGACACTGCCAGGCGT GGCTGCATTGACCAGCGCCggatcttccggatggctcgag-3' and 5'-TGGTTTGTGCTTACAAGGTAAACGAT GCGATTAACATTAATATAACCGTATATATGTAC GCgaagttcctattctctagaaagttaggaacttccatag-3' (80 and

60 nt homology) or 5'-TGCCAGGCGTGGCTGCAT TGACCAGCGCCGgatcttccggatggctcgag-3' and 5'-TAACATTAATATACCGTATATATGTACGCgaagt tctattctctagaaagttaggaacttccatag-3' (29 nt homology). PCR products were purified via Qiagen PCR purification columns before transfection. The sgRNA template primer for PGK was 5'-taatacactactataG TATATATGTACGCTTAGGCGCGtttttagagtag-3'. For tagging of Tubulin 56D, the HR donor primers were 5'-ATGAACGATCTGGTGTCCGAGTACCAGCAGTACCAGGAGGCCACCCCGCAGGAGCAGCCGAGT TCGAGGAGGAGCAGGAGGCTGAGGTTCGACGAG AACggatcttccggatggctcgag-3' and 5'-TACAACCTGAATT CATTGTTCCTCGTTTCATTTTTTTTCGCAAAACA TTGATCGAGAATTCGATTGATTTCCGATTCCGAAT GAAGTTCCTATTCTCTAgaagtaggaacttccatag-3'; the sgRNA template primer was 5'-taatacactactataG ACGAGAATAAATTCGAATgttttagagtag-3'. For tagging of the *blanks* gene, the HR donor primers were 5'-ATCCAAGAAGGCAGCCCGTTACAAACTATC CGCTTTAGTTTGTAAACAAACTATTTGGAACCGAC TATCCACAAAAGgatcttccggatggctcgag-3' and either 5'-GAATGCCGCTAATACGTTTTAAAGTACTACCG GGTGCCAGAAATCATGTGATGTTTTTCGTGAAG TTCCTATTCTCTAgaagtaggaacttccatag-3' (wild-type sequence) or 5'-GAATGCCGCTAATACGTTTTAA AGTACTACCGGGTCCCGCAAGTATGTGATGTT TCGTGAAGTTCCTATTCTCTAgaagtaggaacttccatag-3' (two point mutations in 3'-UTR located CRISPR target site underlined); the sgRNA template primer was 5'-taatacactactatagACGAAAACATCACATGATTCTgttttagagtag-3'.

Primers to verify integration at the PGK locus were 5'-GTGGCCGCCACCAAGAACG-3' (PGK-specific, sense) and 5'-GTAGGTTGAATAGTATATTCCAACAGCATATG-3' (within the *copia*-promoter, antisense). To detect FLP-mediated excision of the resistance cassette, we used primers 5'-ggatcttccggatggctcgag-3' (common linker between target protein and tags, sense) and 5'-CCTCAATGGACAGCAACTTTGCCC-3' (PGK specific, downstream antisense).

All oligonucleotides for cloning and generation of targeting PCR products were ordered from Eurofins/MWG (Ebersberg, Germany) at the appropriate synthesis scale and only standard purification (high purity salt free or HPSF).

The relevant plasmids will be made available through Addgene.

The dsRNA for knock-down was prepared as previously described (71) with PCR primers containing a T7 promoter. For *lig4* RNAi the primers were 5'-taatacactactataggCCCAATGATCCAAAGTGTTTTTGCA-3' and 5'-taatacactactatagGGAAGTAGGATGCCTTCGCA-3', for *exol* RNAi the primers were 5'-taatacactactatagggCCGAAGTCGCTTCTTCGCCAC-3' and 5'-taatacactactatagggGGACAGGTTTCGTCTGGAAGC-3'. The *Renilla* luciferase control construct has been described before (70).

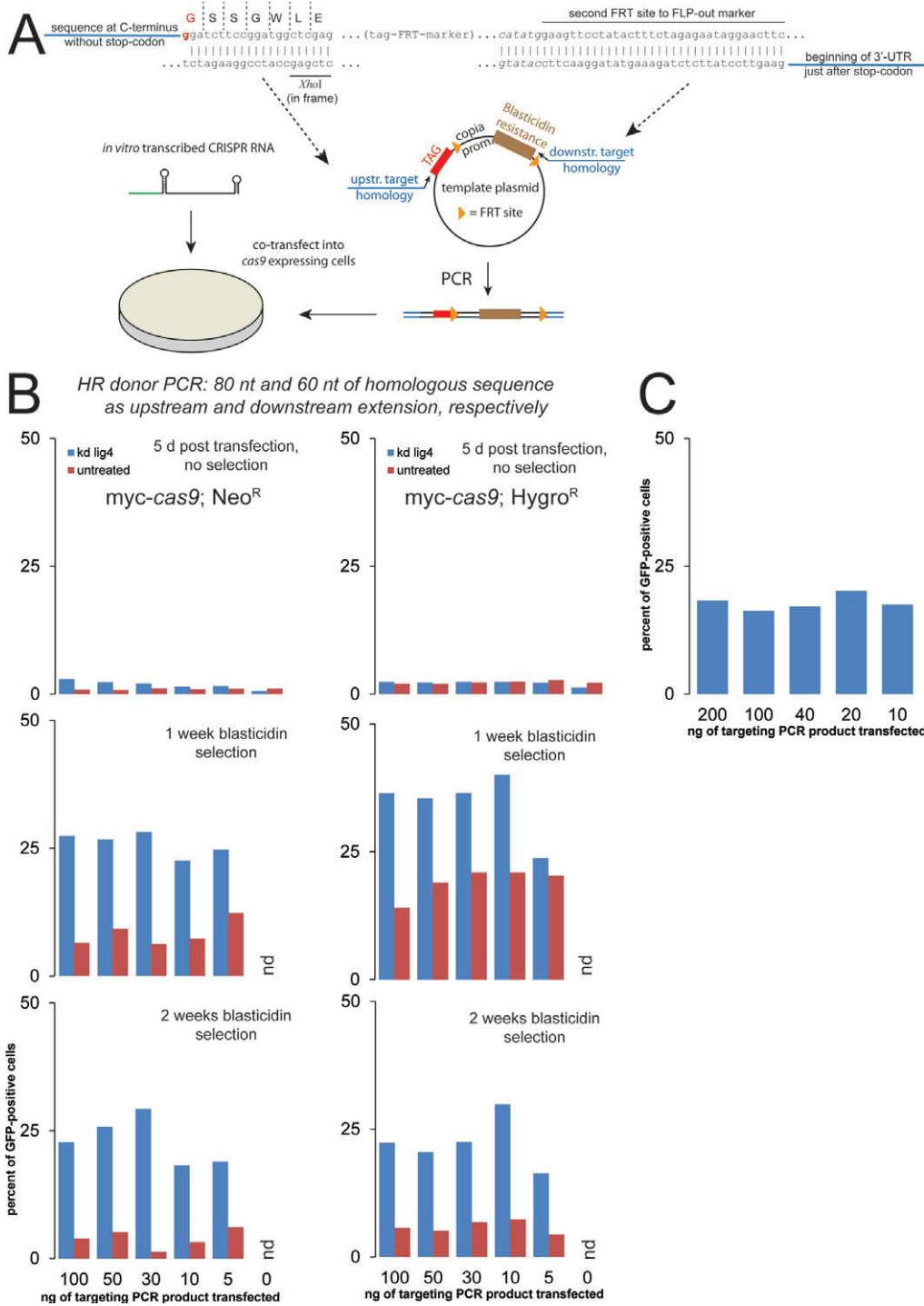


Figure 3. Using CRISPR-mediated induction of DNA double-strand breaks for genome editing via HR. (A) Schematic representation of the targeting construct design and transfection; the length of the homology arms appended via the PCR primers was 80 nt upstream and 60 nt downstream. Further details, including targeting design, are provided in the supplementary information. (B) Assessment of targeting efficiency using the integration of a GFP moiety at the C-terminus of PGK. The three panels represent flow cytometry analysis 5 days after the transfection (top) and after one and two weeks of blasticidin selection (middle and bottom, respectively). The quantity of PCR product transfected (indicated below the bottom panel) refers to the amount applied to one well in a 96-well plate, the amount of sgRNA employed was 100 ng per well. We compared two stable cell clones expressing myc-*cas9* that were independently selected (neomycin and hygromycin resistance) either as untreated cells (red bars) or as cells where the NHEJ factor lig4 had been depleted before transfection (two sequential dsRNA treatments, blue bars). The slightly higher proportion of cells detected as GFP positive after the first split in blasticidin-containing medium may have been caused by a not yet complete antibiotic selection process; non-integrating but transfected cells will be transiently resistant to blasticidin. Since the GFP fluorescence is low (endogenous levels), it is challenging to completely discriminate dying cells from weak GFP positive cells. Note that in the control transfection without any targeting PCR product, no cells could be measured upon selection (nd = none detected). (C) An independent experiment with the neomycin selected myc-*cas9* cell clone yielded comparable results. For this experiment, only one dsRNA treatment to deplete lig4 was performed; this may explain the slightly lower efficiency compared with B.

Cell culture and transfections

Cells were cultured in standard Schneider's medium supplemented with 10% fetal bovine serum (both Bio&Sell, Nürnberg/Germany). Transfections were performed with Fugene HD (Promega) and selection of hygromycin or neomycin resistant cells was done as previously described (72). Selection of blasticidin-resistant cells was achieved at a concentration of 25 $\mu\text{g/ml}$ blasticidin-S (Life Technologies). During selection, cells were split between 1:5 and 1:10 once per week. We employed a Becton Dickinson FACSCalibur flow cytometer equipped with a 96-well plate autosampler for quantification of PGK-GFP expression. Data analysis was performed using flowing software version 2.5.0 (<http://www.flowingsoftware.com/>). RNAi was induced by adding dsRNA to the growth medium at a concentration of 10 $\mu\text{g/ml}$. In our experience, serum withdrawal is not required for efficient RNAi in S2-cells. Four days after induction of RNAi the cells were either transfected or split 1:10 for a second round of RNAi targeting the same gene. Fluorescence microscopy was performed on a Leica TCS SP2 confocal microscope.

RESULTS

The *Streptococcus pyogenes cas9* nuclease efficiently cleaves chromosomal DNA of cultured *Drosophila* cells

Our goal was to introduce a protein tag at the C-terminus with only short regions of homology, therefore selection of suitable CRISPR target sites was very much restrained. We thus tested whether variations of CRISPR RNA length can be tolerated. To this end, we constructed a direct repeat of two GFP coding sequences interrupted by one copy of the *I-SceI* nuclease recognition sequence (Figure 1A). The reporter is conceptually equivalent to comparable systems developed for mammalian cell culture (73). Although in our case the first GFP sequence is complete (while the second one lacks the original ATG start codon), the complete reporter cassette only yields moderate green fluorescence in S2-cells. Upon cleavage by the *I-SceI* nuclease, repair via HR between the two GFP sequences can occur (most likely via the single-strand annealing pathway). This leads to a substantial increase in GFP fluorescence, which can be quantified by flow cytometry in stable cell clones. Apparently, in the complete reporter cassette the long 3'-UTR, comprising the second GFP copy, results in inefficient protein expression. PCR analysis of genomic DNA confirmed rearrangement of the locus (data not shown). Thus, the conversion to high GFP expressing cells in our reporter culture indicates DNA cleavage. As an alternative to HR, the NHEJ pathway may also repair the cleaved reporter. In this case, however, no increase in GFP expression will be observed. By depleting the cellular pool of either DNA ligase 4 (*lig4*), an essential enzyme for NHEJ, or the HR factor exonuclease 1 (*exo1*) (74) via RNA interference, the corresponding pathways can be substantially repressed. Subsequent transfection of the *I-SceI* nuclease and analysis of GFP expression by flow cytometry correctly reflected action of the corresponding repair pathway (Figure 1B). According to preliminary PCR analysis (not shown), it is indeed the repair via HR/single-strand annealing that leads to cells express-

ing high levels of GFP, but a more thorough analysis of the repair mechanism(s) is required and will be presented elsewhere. For the purpose of this manuscript, the important feature of the reporter is that a rearrangement of the locus is quantifiable and triggered only upon DNA cleavage. Thus, our GFP-*I-SceI* seq.-GFP reporter construct can be used to indirectly estimate and compare cleavage efficiencies.

We introduced an expression construct for *cas9* with humanized codon bias (*hcas9*, (47) with an additional myc-tag appended) and selected cell clones expressing both, myc-*hcas9* and the GFP-*I-SceI* seq.-GFP reporter. In these cells, introduction of an *in vitro* transcribed CRISPR RNA with corresponding sequence can direct *cas9*-mediated cleavage of the GFP reporter (Figure 1C). The template for *in vitro* transcription of the CRISPR RNA can be conveniently generated by PCR (see supplementary information for details).

Length variations of the CRISPR sgRNA are tolerated

We used this system to compare cleavage efficiencies between CRISPR RNA sequences with the capacity to form RNA-DNA hybrids between 13 nt and 25 nt in length. The minimal target length construct contained two additional guanines at the 5'-end to ensure efficient transcription by T7 RNA polymerase. Among the RNA-DNA hybrid lengths tested, we found that sequences between 17 nt and 25 nt efficiently programmed *cas9* to cleave the reporter DNA with no negative effects induced by shortening of the targeting sequence length to less than 21 nt; in fact, a trend towards higher cleavage efficiencies with 17 nt and 18 nt targeting length is apparent but should be validated with an independent sequence. The shortest construct, however, did not result in detectable cleavage and repair (Figure 2B). Recent reports have described that an NAG trinucleotide may also serve as PAM for *S. pyogenes cas9*, though cleavage efficiency is only about one-fifth of the one observed with a GG dinucleotide (75-77). When cleavage of our GFP-*I-SceI* seq.-GFP reporter was programmed adjacent to an NAG PAM (see Figure 2A) we observed indeed that recombination-mediated repair is induced, albeit with much lower efficiency (Figure 2C). The 3'-end of the artificial sgRNA scaffold we employ to program *cas9* cleavage specificity has been shortened relative to the naturally occurring CRISPR/tracr-RNA pair (4,47). We thus reasoned that it may be extended in order to append further functional sequences, e.g. a longer stretch of double-stranded RNA or chemical modifications to enhance cellular uptake, if desired. We first prolonged the 3'-end with a short sequence homologous to the reporter sequence downstream of the *I-SceI* cleavage site (we chose the opposite DNA strand). The rationale was to test whether this addition could restore efficient cleavage of our shortest CRISPR targeting RNA construct by providing the potential to form additional base-pairs 3' to the PAM. While cleavage of our reporter occurred with the same efficiency for the 3'-extended, 17-25 nt targeting region CRISPR RNAs, the shortest construct still did not show any detectable cleavage activity (Figure 2D).

Cleavage by the *I-SceI* meganuclease is of comparable efficiency and both systems are most likely limited by the transfection efficiency (we routinely obtain 60-70%). The

RNA–DNA hybrid length can be considerably more variable than what is observed in the original organism *S. pyogenes*; changing its length between 17 nt and 25 nt did not reduce cleavage efficiency. Since both the T7 RNA polymerase promoter we exploited for *in vitro* transcription and the frequently employed U6 snRNA promoter for *in vivo* transcription of the CRISPR RNA initiate most efficiently with a guanosine nucleotide, our observation expands the repertoire of biochemically accessible CRISPR target sites within the genome. This is consistent with mutational studies demonstrating that cleavage specificity is primarily determined by the first 12 nt of the CRISPR targeting portion (counting from the NGG PAM) (4,48,51,75–78). Our observation agrees with previous reports demonstrating that the CRISPR RNA can be extended at its 3'-end (61,78). A contribution of further guideRNA-target DNA base pairs that could rescue the activity of an excessively shortened guide RNA was however not observed with the particular extension we chose. While our manuscript was in revision, Fu et al. (55) reported the use of truncated guide RNAs in human cells. Consistent with our results, guide RNA lengths of 17 and 19 nt showed full cleavage activity while a 15 nt guide RNA appeared to be inactive in their experiments. Furthermore, variations of CRISPR guide RNA length have been tolerated to a certain extent when the CRISPR/*cas9* system was repurposed into a tool for gene regulation (61). Extension of the sgRNA was also tested when Ren et al. used microinjection of DNA templates for sgRNA transcription in *Drosophila* embryos. One of their constructs used the 5'- and 3'-UTR of the *nos* gene to flank the sgRNA; this construct did not yield heritable mutations, though technical aspects such as the low survival rate may in part be responsible for this finding (21).

In summary, our experiments demonstrate that the *S. pyogenes* CRISPR/*cas9* system efficiently cleaves chromosomal DNA in cultured *Drosophila* cells. Our results extend the findings of an independent report (Basset et al. (79)) by a system that features clonally selected cell lines with stable *cas9* expression and that allows for only transient activation of the *cas9* nuclease via transfection of an *in vitro* transcribed guide RNA.

Integration of PCR-based homologous recombination donors is efficient in *Drosophila* S2 cells

Chromosomal DNA double-strand breaks stimulate DNA repair and this can be used for genome editing via HR. A prominent example of this is the addition of a GFP moiety or other protein tags to the endogenous copy of a given protein-coding gene. To this end, the protein tag sequence and a selectable marker are decorated with flanking arms of homologous sequence, typically in the kilobase range (12,14,15,39,41,80–82), thus specifying the desired site of integration. In analogy to elegant approaches used for protein tagging in budding yeast (64), we tested whether these homology arms can be appended via simple additions to PCR primers. This approach is extraordinarily convenient as no cloning steps are involved; the downside is a potentially much lower efficiency of HR due to the length limitations in oligonucleotide synthesis.

We constructed a series of plasmids that serve as templates during PCR to generate a given HR substrate for C-terminal protein tags. The resistance marker is the blasticidin deaminase (BSD) gene from *Aspergillus terreus* allowing for very rapid selection, driven by the *copA* promoter and flanked by FRT sites to permit removal after selection. The sequence coding for a given epitope tag is inserted directly adjacent to the upstream FRT site using *XhoI* and *SpeI* restriction sites. Note that *XhoI/SalI* and *SpeI/XbaI/NheI* have compatible overhangs that can be used for cloning if one of the original sites is present within the desired tag. If the sequence used for primer annealing during PCR is chosen upstream of the *XhoI* site, a series of HR substrates for introducing different tags can be generated with a single set of homology-containing long oligonucleotides. The general setup and resulting linker sequence are detailed in Figure 3A and the supplementary information. Currently, vectors are available with templates for GFP, 2x FLAG and a Twin-STREP-tag (83).

We tested the system by tagging the endogenous copy of PGK with GFP. Since the protein is robustly expressed, the fluorescence induced upon successful tagging can be detected and thus conveniently quantified in a flow cytometer. Previous attempts to use HR for genome editing in *Drosophila* have revealed that inactivating the NHEJ pathway via a mutation in the *lig4* gene can greatly increase the proportion of marker insertions at the desired locus. We therefore compared untreated S2 cells stably expressing the *cas9* nuclease to the same cells with prior depletion of *lig4* via RNA interference. The CRISPR RNA directing cleavage 2 nt 5' to the stop codon in the PGK gene was co-transfected with decreasing amounts of PCR products containing 80 nt and 60 nt of sequence homology upstream and downstream, respectively. We quantified the proportion of GFP positive cells 5 days after transfection, then after one and two weeks of blasticidin selection. Given the limited extent of homologous sequence, we were surprised to see that even in the absence of any selection, up to 2% of the cell population showed detectable GFP expression. This population increased upon blasticidin selection to up to 35%. The amount of co-transfected targeting construct did not appear to be very critical, though a dose of 10 to 30 ng (per well of a 96-well plate) gave the best results. These values were obtained with cells that had been pre-treated with dsRNA to deplete *lig4*. In untreated cells, no more than 6% of the blasticidin-resistant cells were GFP positive after two weeks of selection, demonstrating that NHEJ mediated off-target integration is responsible for at least a portion of the blasticidin-resistant, non-fluorescent cells (Figure 3B). The reproducibility of this success rate is good as similar results were obtained in a parallel experiment with an independent clone of *cas9* expressing cells (Figure 3B) and a fully independent biological replicate (Figure 3C). Correct and full-length integration was verified by genomic PCR (Figure 4A), excision of the blasticidin-resistance cassette upon transfection of a FLP-recombinase expression plasmid (Figure 4B) and western blotting to confirm the predicted size of GFP-, Flag- and Strep-reactive bands (Figure 4C).

There are two possibilities why we derive a substantial amount of blasticidin-resistant, GFP-negative cells. Either

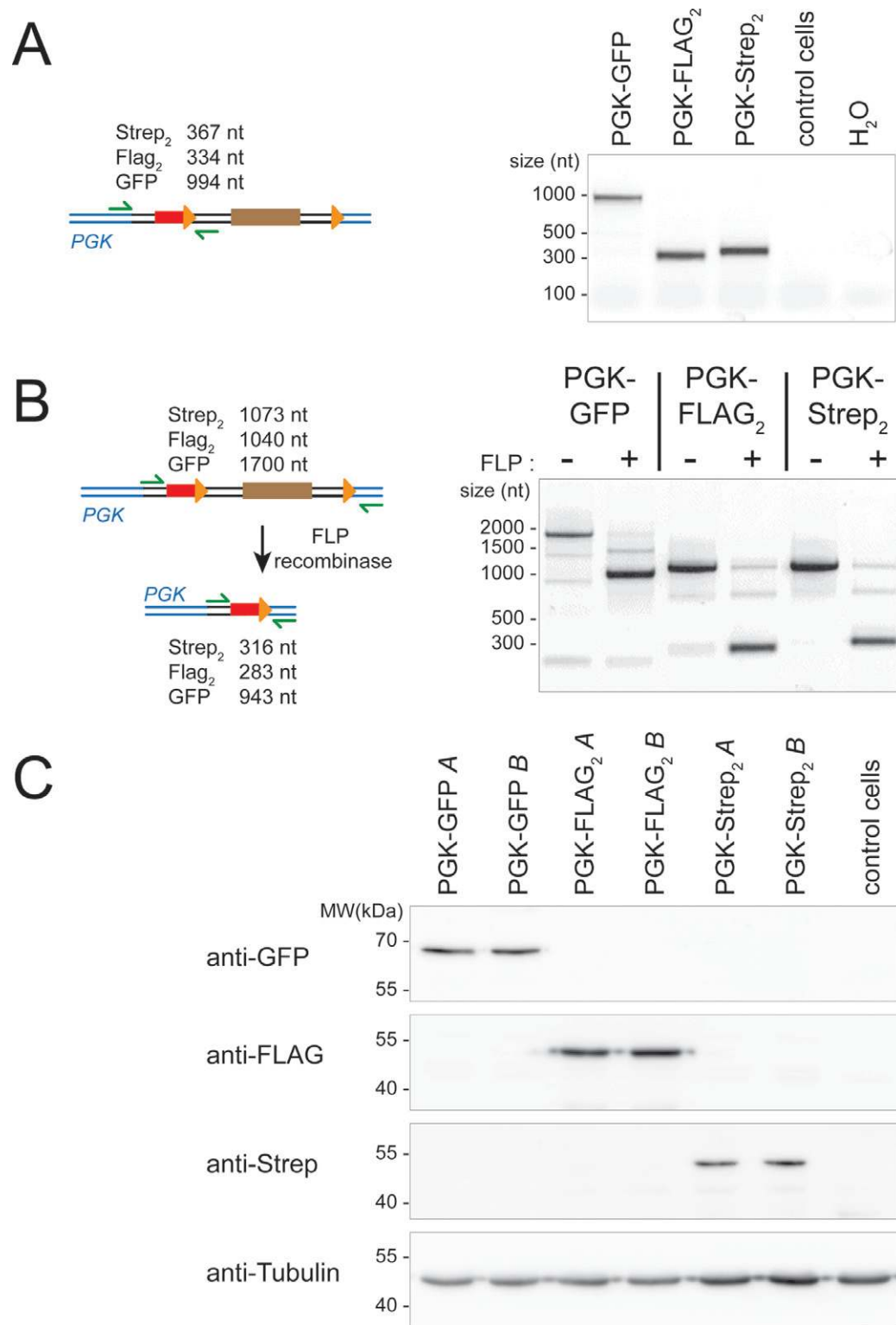


Figure 4. Molecular characterization of epitope tags introduced at the PGK locus. **(A)** PCR analysis to verify integration at the PGK locus. The primer locations and expected product sizes are indicated on the left, the specific upstream and downstream primers were chosen outside of the homology-region used for targeting. The PCR products were sequenced to verify their identity. **(B)** FLP-mediated excision of the resistance cassette. The primer locations and expected product sizes before and after excision are indicated on the left. FLP-mediated excision was further verified by sequencing the PCR products obtained for all three tags after introduction of FLP recombinase. **(C)** Western blot to verify that a defined protein of the expected size is tagged using our protocol. The blots were probed with monoclonal antibodies against GFP (clone B-2 obtained from Santa Cruz), the FLAG-tag (clone M2 obtained from Sigma) and the Strep-tag (Strep-Mab HRP, obtained from IBA in Göttingen/Germany). Detection of tubulin served as a loading control (below, clone E7 obtained from the Developmental Studies Hybridoma Bank). For each tag, two parallel technical replicates were analyzed (= parallel transfection and selection, referred to as A and B).

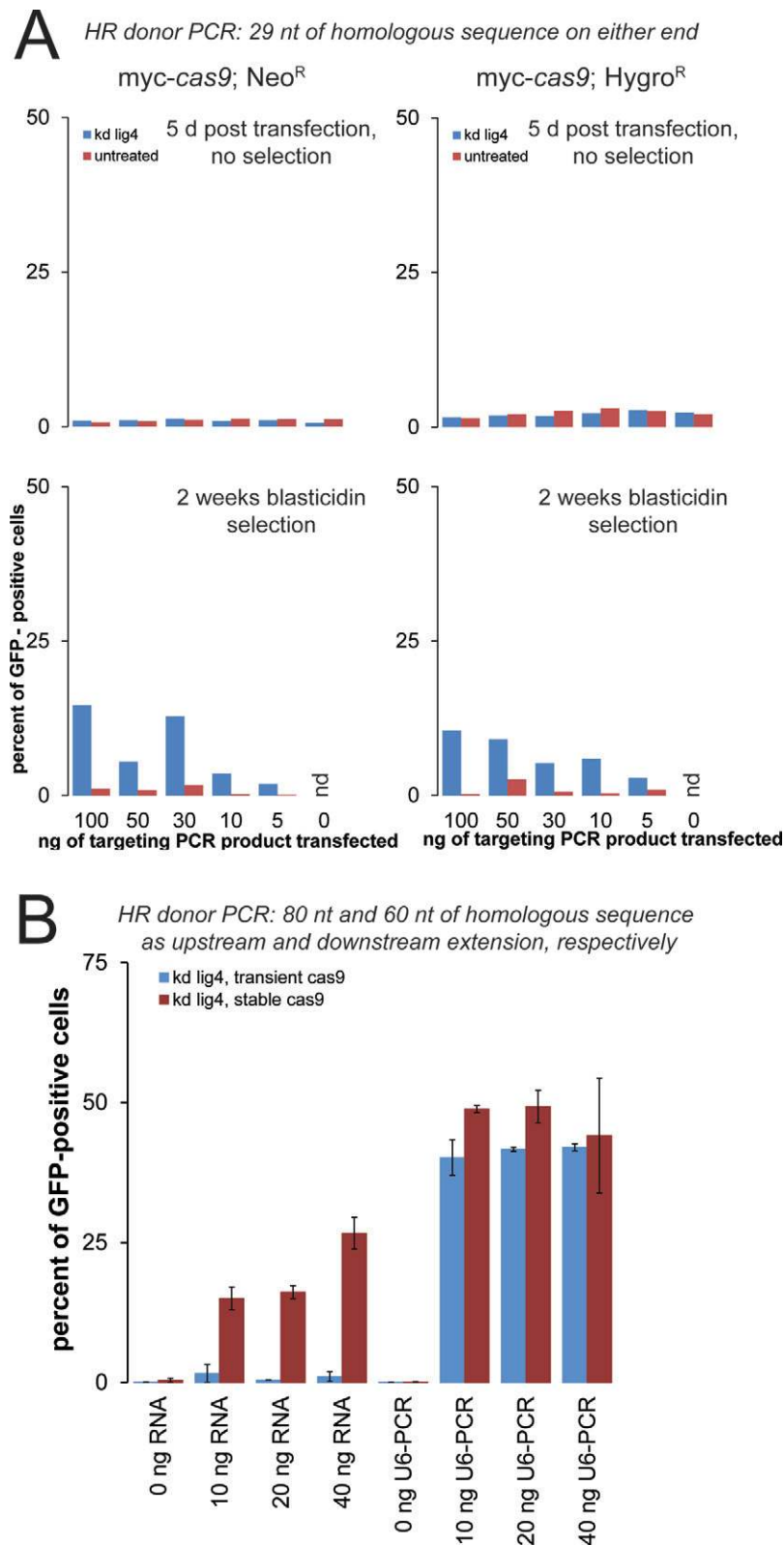


Figure 5. Short homology arms can direct site-specific integration and increased targeting efficiency with U6-promoter based sgRNA expression. (A) We performed an analogous experiment to Figure 3B using PCR primers with only 29 nt of sequence homology at either end. Although the efficiency is clearly lower than what was observed with longer homology, the short arms are able to direct site-specific integration of the PCR product. (B) Transfection of U6-promoter sgRNA template fusions can further boost the recovery of tag-expressing cells. The amount of sgRNA (obtained by *in vitro* transcription) or U6-promoter-sgRNA template (obtained by overlap extension PCR) is indicated below the diagram. Note that for the *in vitro* transcribed sgRNA, the maximal amount used (40 ng) is less than in Figure 3 (100 ng). Blue bars depict the results from transient *cas9* expression, while red bars depict the results obtained with a stable *cas9* expressing clone (hygromycin-resistance). The bars represent the mean of two fully independent biological replicates, error bars indicate the range of the individual values.

the PCR cassette spontaneously integrated elsewhere in the genome in spite of our *lig4* depletion, or small mistakes during HR or oligonucleotide synthesis have introduced a shift in the reading frame and thus prevented GFP expression. To distinguish these possibilities, we sequenced the PCR products from Figure 2A directly (i.e. without cloning and selection of individual colonies). The sequencing traces showed only minor signs of sequence heterogeneity, indicating that frameshift mutations are unlikely to account for many of the GFP-negative cells. Off-target integration of the tagging-cassette is therefore a frequent event in the selected population and could—at least in theory—lead to expression of the protein tag on unintended proteins. So far, we have not detected any corresponding signal in our Western blots or in the fluorescence microscopy results (see below, Figure 6). Furthermore, transfection of the GFP tagging cassette in the absence of a corresponding DNA break does not yield significant amounts of GFP-positive cells (see below, Figure 5). The non-expressing cells may have suffered from the mutagenic insertion of the tagging cassette into off-target locations. However, since these cells do not appear to express the tag, this should not compromise conclusions based on detection of the tag.

Is it possible to further shorten the homology arms contributed by the PCR primers? We performed the same experiment with only 29 nt of upstream and downstream homology. Without selection, we only detected background levels of GFP fluorescence, but upon blasticidin treatment up to 15% of the selected cell population became GFP positive. Again, the depletion of *lig4* resulted in a clear increase in GFP-positive cells. The success of this experiment demonstrates that flanking homology as short as 29 nt can suffice to direct site-specific integration at an engineered DNA double-strand break (Figure 5A). We nonetheless recommend using longer stretches of homology as this increases the frequency of targeted integration.

Improved efficiency using U6-promoter based sgRNA expression

Since 5'-extensions of the guide RNA did not impair cleavage efficiency (Figure 2), we tested whether the T7-promoter sequence contained in our sgRNA *in vitro* transcription templates can be used to append a U6 promoter via overlap extension PCR. To this end, we amplified the *Drosophila* U6-C promoter with the T7 RNA polymerase promoter sequence as an extension after the annotated transcription start site. Upon overlap-extension PCR, an U6-promoter-sgRNA template fusion product of ~600 nt can be obtained (not shown). We then transfected this construct along with our PGK-GFP targeting PCR product (harboring the longer homology arms) into *lig4*-depleted cells stably expressing the *cas9* nuclease. For a side-by-side comparison, we also included *in vitro* transcribed RNA. In this experiment, we varied the amount of CRISPR guide RNA (or U6-promoter fusion) while keeping the amount of tagging construct constant. Even without *cas9* programming, surviving cells could be recovered after two rounds of blasticidin selection. However, essentially none of these cells showed green fluorescence, further substantiating that the fluorescence we observed is due to on-target integra-

tion. Upon activation of DNA cleavage at the PGK locus, we saw that the U6-promoter fusion constructs resulted in even higher proportions of cells with PGK-GFP fusions than transfection of *in vitro* transcribed RNA (Figure 5B). We also combined transient transfection of a *cas9* expression plasmid, the U6-promoter-gRNA template fusion PCR and the PGK-GFP tagging PCR product into *lig4*-depleted, normal S2-cells. This combination also produced PGK-GFP expressing cells albeit with a slightly reduced efficiency. Co-transfection of *in vitro* transcribed gRNA instead of the U6-promoter fusions was clearly less efficient when combined with transient *cas9* expression, presumably because much of the gRNA template is degraded before the *cas9* protein becomes expressed (Figure 5B).

Application to other genes and introduction of point mutations in HR donors

To confirm the general applicability of our technique, we also designed gRNA and HR donor primers to tag the *β -tubulin 56D* gene. Successful integration of the tagging cassette was confirmed by PCR and western blotting (not shown). Upon introduction of a GFP-tag, cytoskeletal structures could be observed by fluorescence microscopy in interphase cells (Figure 6A, upper row) and the spindle was clearly GFP-labeled in mitotic cells (Figure 6A, bottom row). In contrast, microscopic analysis of PGK-GFP expressing cells revealed a diffuse cytoplasmic and nuclear distribution of GFP fluorescence (Figure 6B). This localization pattern is consistent with results obtained in mammalian cells, where cytoplasmic and nuclear localization of glycolytic enzymes—including PKG—has been observed (reviewed in (84)).

Serial dilution of the blasticidin-resistant cells can be performed to derive cell populations that show uniform expression of the tag. We found this to be straightforward from a starting population where 30% of the cells have the tag integrated at the correct position (see Supplementary Figure S1 for an example clone). This may even yield cells where all chromosomal alleles have been modified (i.e. no untagged protein remains), provided that the protein remains functional or is non-essential.

Stimulation of HR by *cas9*-mediated DNA cleavage depends on the availability of a correspondingly located GG PAM. Ideally, the *cas9*/CRISPR target site is disrupted by integration of the tagging cassette such that neither the HR donor nor the modified genomic locus can be cleaved (see Supplementary Figure S2A for targeting design of *PGK* and *Tub56D* locus). If this ideal design is not possible, cleavage of the donor and modified locus can be prevented by including point mutations (translationally silent if located within the coding sequence) in the flanking homology arms of the PCR primers. On the other hand, this may have adverse effects on the efficiency of HR, in particular because our approach provides only short homology arms. Using the nuclear RNAi-factor *blanks* as target protein, we compared the tagging efficiency for point mutant and wild-type HR donor PCR products for a non-disrupted CRISPR target site. Western Blot Analysis demonstrated that integration of a FLAG-tag was clearly more efficient if the HR donor contained the point mutations (Supplementary Fig-

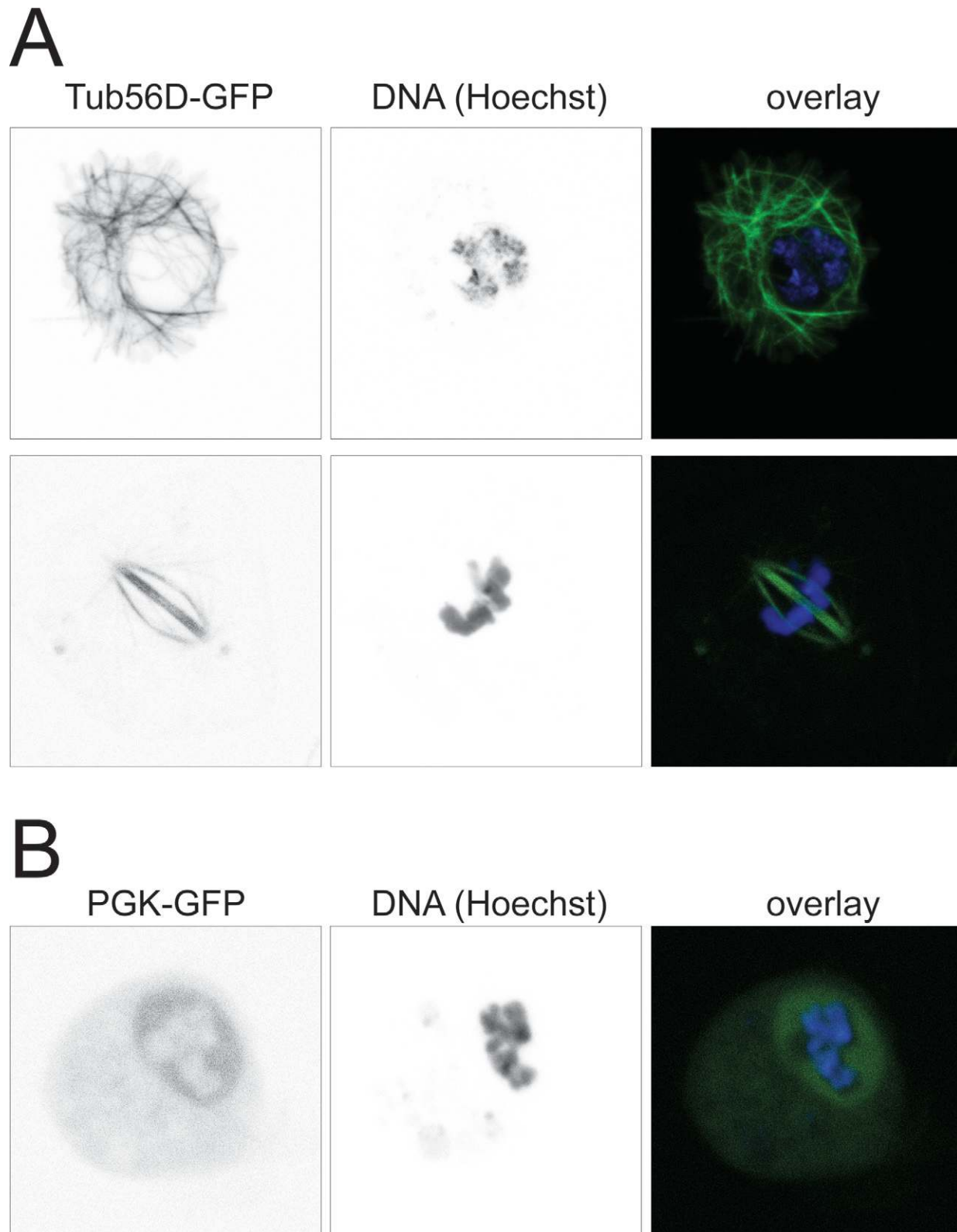


Figure 6. Fluorescence microscopy of Tub56D-GFP and PGK-GFP. **(A)** Tub56D-GFP tagging reveals cytoskeletal structures in interphase cells (top row) and the spindle in mitotic cells (bottom row). **(B)** The GFP-tagged PGK protein is localized to both nucleus and cytoplasm. Live cells were stained with Hoechst 33342, a cell-permeable DNA stain with DAPI-like fluorescence, and imaged on a Leica SP2 confocal microscope. We acquired z-stacks covering the entire thickness of the cells, then calculated the corresponding 2D average projections that are displayed in this figure with the Leica software accompanying the microscope. These high magnification images are representative of the GFP-expressing cells in our cultures. In particular, we did not observe any cells with tubulin-like GFP distribution in the PGK-GFP cultures and *vice versa*.

ure S2B). Thus, the *cas9*/CRISPR mediated cleavage may occur a short distance away from the desired integration site, provided that point mutations can be introduced in the HR donor. Further experiments are required to delimit the maximal distance of DNA cleavage from the desired integration site that still results in highly efficient tagging.

DISCUSSION

Limitations and further applications of the chromosomal tagging approach

The system we are describing here allows for straightforward generation of tagged proteins at their chromosomal locus. This circumvents well-known problems that are inherent to approaches based on plasmid cloning and transient transfection, such as overexpression at heterogeneous levels and the presence of untagged, endogenous protein. However, some general limitations of protein tags remain. First of all, addition of a tag may interfere with the function of the host protein. With plasmid transfections, this can sometimes be masked by the remaining pool of untagged endogenous protein; functionality of the chromosomally tagged protein should therefore be assessed in cell clones where all chromosomal alleles have been tagged. Furthermore, appending a relatively stable protein such as GFP may influence the host protein's stability, leading to altered expression levels despite an unaltered gene dose.

Some fly genes show heterogeneity at the 3'-end due to splice variants that involve the region coding for the protein C-terminus. Fortunately, these cases are the exception rather than the rule. Here, introduction of a C-terminal tag is not only impractical but in addition could influence the outcome of alternative splicing decisions. It may be possible to introduce tags also internally with our method; however, the current template modules all contain a stop codon at the end of the tag and the FRT sequence, which remains after removal of the resistance cassette, will also introduce stop codons depending on the reading frame.

Off-target cleavage by the *cas9* nuclease must be a concern. We have thus far employed CRISPR target sites that are very close to the stop codon of the host protein. Due to the limited sequence space available for DNA cleavage, the choice of CRISPR sites with minimal off-target cleavage potential is essentially impossible in this situation. Further experiments should test whether the DNA cut may be set at a certain distance to the desired integration site, allowing for at least a minimal choice among potential CRISPR target sites. Finally, the use of truncated guide RNAs is not only efficient (see Figure 2) but has also reduced the off-target cleavage rates in cultured human cells (55).

It is also important to consider when delivery of the guide RNA as a transcript or as a U6-promoter-based DNA template is more appropriate. Recent structural analysis has revealed that RNA-binding is required for conformational activation of the *cas9* nuclease (85,86). This likely explains why expression of apo-*cas9* is non-toxic in cells and transgenic flies; delivery of an *in vitro* transcribed guide RNA therefore only produces a transient pulse of CRISPR/*cas9* cleavage activity. If DNA-based guide RNA templates are employed, the potential for spontaneous integration of this

cassette into the genome may lead to permanent *cas9* activation and thus a high rate of off-target cleavage events. We suggest using *in vitro* transcribed guide RNAs and clonal selection of tagged cells if these will represent the basis for a larger set of detailed analyses. In contrast, the U6-promoter-guide RNA template fusions are valuable tools for automated high throughput screening approaches.

The combination of *lig4* RNAi, *cas9*-mediated DNA cleavage, introduction of a PCR-generated HR donor and blasticidin selection can routinely yield cell populations where more than 30% of the cells have a correctly targeted insertion at the desired genomic locus. This is almost within the range of what can be achieved by transient transfections, yet it has the advantage that the endogenous promoter and gene dose are conserved thus avoiding overexpression artifacts. It may be possible to improve the yield of correctly inserted fragments further by optimizing the *lig4* knock-down procedure and/or combining it with knockdown of other NHEJ factors, in particular those responsible for a *lig4*-independent NHEJ-like pathway (87–91).

The tagging at endogenous expression levels evidently limits the scope of our technique to the set of proteins normally present in the S2 cells. Analogous template modules for N-terminal tagging can be devised; they should provide a promoter to force expression of the fusion protein until the selection marker is removed by FLP recombinase. Before marker excision, expression of the targeted genes will thus be constitutive or—depending on the promoter—conditional. Although e.g. interactome studies with a protein that is normally absent from the cellular context are likely of limited value, mis-expression strategies have been successful to identify novel components of e.g. the *Wnt* signalling pathway in flies (92). With a properly engineered readout, genome-scale overexpression may thus represent a valid and feasible screening approach in S2-cells. Our system is also well suited for knock-in of e.g. a luciferase reporter either fused to or in place of the host gene. This permits convenient measurement of gene regulation within the chromosomal context, including the 3'-UTR if FLP-mediated excision of the resistance cassette has been applied. Combining these reporters with genome-wide RNAi screens is then straightforward. In particular, since transient *cas9* expression can suffice to introduce the tags (see Figure 5B), it should be possible to extend the application to other *Drosophila* cell lines, which can also be derived *de novo* from flies with a desired genotype (93).

Are PCR-based HR donors less efficient than the established, plasmid-based approaches using long homologues? We have not attempted a side-by-side evaluation and quantitative comparisons between flies and S2-cells are not reasonable. An earlier attempt to detect genome modification of KC167/M3 cells via HR by the Cherbas lab employed long homology arms, but could not induce a site-specific DNA break (94). It is thus not suitable for direct comparison with our data. The independent work of Bassett et al. (79) in S2-cells demonstrated tagging efficiencies for the *Ago1* gene of around 1% without selection using a 1 kb homology arm. This is in fact roughly comparable with the efficiencies we observed at the PGK locus prior to selection and without depletion of *lig4* (see Figure 3B). Subsequently, the approach by Bassett et al. only removes

the non-transfected cells, increasing the tagging frequency to 4–5% of the selected cells. Our protocol selects for stable integration of a marker and we thus obtained a much higher proportion of cells expressing tagged proteins. Several publications have demonstrated that linear HR donors (66) or circular HR donors with homology arms of less than 200 nt in length are inefficient when micro-injected into fly embryos (63). In contrast, work in cultured mammalian cells demonstrated comparable integration efficiency for long, cloned and short, PCR-based homology arms (95). It therefore remains to be tested whether our PCR-based HR donors can be used efficiently *in vivo* as well.

In addition to the two supplementary figures mentioned in the text, the manuscript is accompanied by a detailed step-by-step protocol for our procedure that includes a collection of several figures giving technical detail on the constructs presented and information for primer design.

AVAILABILITY

The plasmids described in this manuscript will be available through Addgene.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The anti-tubulin antibody clone E7 was obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biology, Iowa City, IA 52242, USA.

FUNDING

Deutsche Forschungsgemeinschaft [SFB646]; Alexander von Humboldt-Professorship from Bundesministerium für Bildung und Forschung [U.G.].

Conflict of interest statement. None declared.

REFERENCES

- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A. and Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, **315**, 1709–1712.
- Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V. and van der Oost, J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, **321**, 960–964.
- Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H. and Moineau, S. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, **468**, 67–71.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
- Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J. and Church, G.M. (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods*, **10**, 1116–1121.
- Gasiunas, G., Barrangou, R., Horvath, P. and Siksnys, V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, E2579–E2586.
- 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.
- Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiacovo, M.P., Church, G.M. and Calarco, J.A. (2013) Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods*, **10**, 741–743.
- Lo, T.W., Pickle, C.S., Lin, S., Ralston, E.J., Gurling, M., Schartner, C.M., Bian, Q., Doudna, J.A. and Meyer, B.J. (2013) Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. *Genetics*, **195**, 331–348.
- Cho, S.W., Lee, J., Carroll, D. and Kim, J.S. (2013) Heritable gene knockout in *Caenorhabditis elegans* by direct injection of Cas9-sgRNA ribonucleoproteins. *Genetics*, **195**, 1177–1180.
- Katic, I. and Grosshans, H. (2013) Targeted heritable mutation and gene conversion by Cas9-CRISPR in *Caenorhabditis elegans*. *Genetics*, **195**, 1173–1176.
- Tzur, Y.B., Friedland, A.E., Nadarajan, S., Church, G.M., Calarco, J.A. and Colaiacovo, M.P. (2013) Heritable *Caenorhabditis elegans* via a CRISPR-Cas9 system. *Genetics*, **195**, 1181–1185.
- Waijers, S., Portegijs, V., Kerver, J., Lemmens, B.B., Tijsterman, M., van den Heuvel, S. and Boxem, M. (2013) CRISPR/Cas9-targeted mutagenesis in *Caenorhabditis elegans*. *Genetics*, **195**, 1187–1191.
- Dickinson, D.J., Ward, J.D., Reiner, D.J. and Goldstein, B. (2013) Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nat. Methods*, **10**, 1028–1034.
- Chen, C., Fenk, L.A. and de Bono, M. (2013) Efficient genome editing in *Caenorhabditis elegans* by CRISPR-targeted homologous recombination. *Nucleic Acids Res.*, **41**, e193.
- 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.
- 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.
- Yu, Z., Ren, M., Wang, Z., Zhang, B., Rong, Y.S., Jiao, R. and Gao, G. (2013) Highly efficient genome modifications mediated by CRISPR/Cas9 in *Drosophila*. *Genetics*, **195**, 289–291.
- Kondo, S. and Ueda, R. (2013) Highly improved gene targeting by germline-specific cas9 expression in *Drosophila*. *Genetics*, **195**, 715–721.
- Sebo, Z.L., Lee, H.B., Peng, Y. and Guo, Y. (2013) A simplified and efficient germline-specific CRISPR/Cas9 system for *Drosophila* genomic engineering. *Fly (Austin)*, **8**, 52–57.
- Ren, X., Sun, J., Housden, B.E., Hu, Y., Roesel, C., Lin, S., Liu, L.P., Yang, Z., Mao, D., Sun, L. *et al.* (2013) Optimized gene editing technology for *Drosophila melanogaster* using germ line-specific Cas9. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 19012–19017.
- Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. and Sheen, J. (2013) Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.*, **31**, 688–691.
- Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D. and Kamoun, S. (2013) Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.*, **31**, 691–693.
- Xie, K. and Yang, Y. (2013) RNA-guided genome editing in plants using a CRISPR-Cas system. *Mol. Plant*, **6**, 1975–1983.
- Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B. and Weeks, D.P. (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Res.*, **41**, e188.
- Upadhyay, S.K., Kumar, J., Alok, A. and Tuli, R. (2013) RNA guided genome editing for target gene mutations in wheat. *G3 (Bethesda)*, **3**, 2233–2238.
- 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.
- Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Kaini, P., Sander, J.D., Joung, J.K., Peterson, R.T. and Yeh, J.R. (2013) Heritable and precise

- zebrafish genome editing using a CRISPR-Cas system. *PLoS ONE*, **8**, e68708.
29. Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X. *et al.* (2013) Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.*, **41**, e141.
 30. Jao, L.E., Wente, S.R. and Chen, W. (2013) Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 13904–13909.
 31. Auer, T.O., Duroure, K., De Cian, A., Concordet, J.P. and Del Bene, F. (2014) Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res.*, **24**, 142–153.
 32. Sung, Y.H., Kim, J.M., Kim, H.T., Lee, J., Jeon, J., Jin, Y., Choi, J.H., Ban, Y.H., Ha, S.J., Kim, C.H. *et al.* (2014) Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. *Genome Res.*, **24**, 125–131.
 33. Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C. and Schmid, B. (2013) Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development*, **140**, 4982–4987.
 34. Nakayama, T., Fish, M.B., Fisher, M., Oomen-Hajagos, J., Thomsen, G.H. and Grainger, R.M. (2013) Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in *Xenopus tropicalis*. *Genesis*, **51**, 835–843.
 35. Shen, B., Zhang, J., Wu, H., Wang, J., Ma, K., Li, Z., Zhang, X., Zhang, P. and Huang, X. (2013) Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res.*, **23**, 720–723.
 36. Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F. and Jaenisch, R. (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, **153**, 910–918.
 37. Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L. and Jaenisch, R. (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*, **154**, 1370–1379.
 38. Fujii, W., Kawasaki, K., Sugiura, K. and Naito, K. (2013) Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res.*, **41**, e187.
 39. Zhou, J., Shen, B., Zhang, W., Wang, J., Yang, J., Chen, L., Zhang, N., Zhu, K., Xu, J., Hu, B. *et al.* (2014) One-step generation of different immunodeficient mice with multiple gene modifications by CRISPR/Cas9 mediated genome engineering. *Int. J. Biochem. Cell Biol.*, **46**, 49–55.
 40. Mashiko, D., Fujihara, Y., Satouh, Y., Miyata, H., Isotani, A. and Ikawa, M. (2013) Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. *Sci. Rep.*, **3**, 3355.
 41. Gennequin, B., Otte, D.M. and Zimmer, A. (2013) CRISPR/Cas-induced double-strand breaks boost the frequency of gene replacements for humanizing the mouse *Cnr2* gene. *Biochem. Biophys. Res. Commun.*, **441**, 815–819.
 42. Hu, X., Chang, N., Wang, X., Zhou, F., Zhou, X., Zhu, X. and Xiong, J.W. (2013) Heritable gene-targeting with gRNA/Cas9 in rats. *Cell Res.*, **23**, 1322–1325.
 43. Yang, D., Xu, J., Zhu, T., Fan, J., Lai, L., Zhang, J. and Chen, Y.E. (2014) Effective gene targeting in rabbits using RNA-guided Cas9 nucleases. *J. Mol. Cell Biol.*, **6**, 97–99.
 44. Tan, W., Carlson, D.F., Lancto, C.A., Garbe, J.R., Webster, D.A., Hackett, P.B. and Fahrenkrug, S.C. (2013) Efficient nonmeiotic allele introgression in livestock using custom endonucleases. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 16526–16531.
 45. Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W. *et al.* (2014) Generation of gene-modified *Cynomolgus* Monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell*, **156**, 836–843.
 46. Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. and Doudna, J. (2013) RNA-programmed genome editing in human cells. *Elife*, **2**, e00471.
 47. 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.
 48. 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.
 49. 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.
 50. Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.F., Sontheimer, E.J. and Thomson, J.A. (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 15644–15649.
 51. Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S. and Kim, J.S. (2013) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.*, **24**, 132–141.
 52. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. and Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nature Protoc.*, **8**, 2281–2308.
 53. 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.* (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, **154**, 1380–1389.
 54. Cradick, T.J., Fine, E.J., Antico, C.J. and Bao, G. (2013) CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.*, **41**, 9584–9592.
 55. Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. and Jung, J.K. (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.*, **32**, 279–284.
 56. Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S. and Qi, L.S. (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.*, **8**, 2180–2196.
 57. Mali, P., Esvelt, K.M. and Church, G.M. (2013) Cas9 as a versatile tool for engineering biology. *Nat. Methods*, **10**, 957–963.
 58. Cheng, A.W., Wang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B. and Jaenisch, R. (2013) Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.*, **23**, 1163–1171.
 59. Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W. *et al.* (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods*, **10**, 973–976.
 60. Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F. and Marraffini, L.A. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.*, **41**, 7429–7437.
 61. Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. and Lim, W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173–1183.
 62. Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S. *et al.* (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, **155**, 1479–1491.
 63. Beumer, K.J. and Carroll, D. (2014) Targeted genome engineering techniques in *Drosophila*. *Methods*, doi: 10.1016/j.ymeth.2013.12.002.
 64. Longtine, M.S., McKenzie, A. 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P. and Pringle, J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast*, **14**, 953–961.
 65. Bozas, A., Beumer, K.J., Trautman, J.K. and Carroll, D. (2009) Genetic analysis of zinc-finger nuclease-induced gene targeting in *Drosophila*. *Genetics*, **182**, 641–651.
 66. 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. U.S.A.*, **105**, 19821–19826.
 67. Forstemann, K., Horwich, M.D., Wee, L., Tomari, Y. and Zamore, P.D. (2007) *Drosophila* microRNAs are sorted into functionally distinct argonaute complexes after production by Dicer-1. *Cell*, **130**, 287–297.
 68. Rong, Y.S. and Golic, K.G. (2000) Gene targeting by homologous recombination in *Drosophila*. *Science*, **288**, 2013–2018.
 69. Hartig, J.V. and Forstemann, K. (2011) Loqs-PD and R2D2 define independent pathways for RISC generation in *Drosophila*. *Nucleic Acids Res.*, **39**, 3836–3851.

70. Michalik, K.M., Bottcher, R. and Forstemann, K. (2012) A small RNA response at DNA ends in *Drosophila*. *Nucleic Acids Res.*, **40**, 9596–9603.
71. Hartig, J.V., Esslinger, S., Bottcher, R., Saito, K. and Forstemann, K. (2009) Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences. *EMBO J.*, **28**, 2932–2944.
72. Shah, C. and Forstemann, K. (2008) Monitoring miRNA-mediated silencing in *Drosophila melanogaster* S2-cells. *Biochim. Biophys. Acta*, **1779**, 766–772.
73. Pierce, A.J., Johnson, R.D., Thompson, L.H. and Jasin, M. (1999) XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev.*, **13**, 2633–2638.
74. Garcia, V., Phelps, S.E., Gray, S. and Neale, M.J. (2011) Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. *Nature*, **479**, 241–244.
75. Jiang, W., Bikard, D., Cox, D., Zhang, F. and Marraffini, L.A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.*, **31**, 233–239.
76. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.*, **31**, 827–832.
77. Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A. and Liu, D.R. (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.*, **31**, 839–843.
78. Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L. and Church, G.M. (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.*, **31**, 833–838.
79. Bassett, A.R. and Liu, J.L. (2014) CRISPR/Cas9 and Genome Editing in *Drosophila*. *J. Genet. Genomics*, **41**, 7–19.
80. Baena-Lopez, L.A., Alexandre, C., Mitchell, A., Pasakarnis, L. and Vincent, J.P. (2013) Accelerated homologous recombination and subsequent genome modification in *Drosophila*. *Development*, **140**, 4818–4825.
81. 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. U.S.A.*, **106**, 8284–8289.
82. Gratz, S.J., Ukken, F.P., Rubinstein, C.D., Thiede, G., Donohue, L.K., Cummings, A.M. and O'Connor-Giles, K.M. (2014) Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics*, doi: 10.1534/genetics.113.160713.
83. Schmidt, T.G., Batz, L., Bonet, L., Carl, U., Holzapfel, G., Kiem, K., Matulewicz, K., Niermeier, D., Schuchardt, I. and Stanar, K. (2013) Development of the Twin-Strep-tag(R) and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr. Purif.*, **92**, 54–61.
84. Ronai, Z. (1993) Glycolytic enzymes as DNA binding proteins. *Int. J. Biochem.*, **25**, 1073–1076.
85. Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S. *et al.* (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, **343**, 1247997, doi: 10.1126/science.1247997.
86. Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F. and Nureki, O. (2014) Crystal structure of cas9 in complex with guide RNA and target DNA. *Cell*, **156**, 935–949.
87. White, T.B. and Lambowitz, A.M. (2012) The retrohoming of linear group II intron RNAs in *Drosophila melanogaster* occurs by both DNA ligase 4-dependent and -independent mechanisms. *PLoS Genet.*, **8**, e1002534.
88. Garcia, A.M., Salomon, R.N., Witsell, A., Liepkalns, J., Calder, R.B., Lee, M., Lundell, M., Vijg, J. and McVey, M. (2011) Loss of the bloom syndrome helicase increases DNA ligase 4-independent genome rearrangements and tumorigenesis in aging *Drosophila*. *Genome Biol.*, **12**, R12.
89. Yu, A.M. and McVey, M. (2010) Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res.*, **38**, 5706–5717.
90. Chan, S.H., Yu, A.M. and McVey, M. (2010) Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in *Drosophila*. *PLoS Genet.*, **6**, e1001005.
91. McVey, M., Radut, D. and Sekelsky, J.J. (2004) End-joining repair of double-strand breaks in *Drosophila melanogaster* is largely DNA ligase IV independent. *Genetics*, **168**, 2067–2076.
92. Schertel, C., Huang, D., Bjorklund, M., Bischof, J., Yin, D., Li, R., Wu, Y., Zeng, R., Wu, J., Taipale, J. *et al.* (2013) Systematic screening of a *Drosophila* ORF library in vivo uncovers Wnt/Wg pathway components. *Dev. Cell*, **25**, 207–219.
93. Simcox, A., Mitra, S., Truesdell, S., Paul, L., Chen, T., Butchar, J.P. and Justiniano, S. (2008) Efficient genetic method for establishing *Drosophila* cell lines unlocks the potential to create lines of specific genotypes. *PLoS Genet.*, **4**, e1000142.
94. Cherbas, L. and Cherbas, P. (1997) “Parahomologous” gene targeting in *Drosophila* cells: an efficient, homology-dependent pathway of illegitimate recombination near a target site. *Genetics*, **145**, 349–358.
95. Orlando, S.J., Santiago, Y., DeKolver, R.C., Freyvert, Y., Boydston, E.A., Moehle, E.A., Choi, V.M., Gopalan, S.M., Lou, J.F., Li, J. *et al.* (2010) Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic acids res.*, **38**, e152.