Highly frequent undesired insertional mutagenesis during *Drosophila* genome editing

Emma Källstig*[1,2], Evelyne Ruchti*[1], Medha Raman[1], Jamshid Asadzadeh[1], Bernard L. Schneider[1,2], and Brian D. McCabe[1]✉

[1]Brain Mind Institute, [2]Bertarelli Platform for Gene Therapy, EPFL - Swiss Federal Institute of Technology, Lausanne, Switzerland

Summary

CRISPR/Cas9 based genome editing employing Homology Directed Repair (HDR) from template vector sequences is a widely used technique to enable precise insertions, deletions or modifications to genes. Here, we describe an undesired and highly frequent editing event when using conventional CRISPR/Cas9 plus HDR methods for *Drosophila melanogaster* germline genome editing. We find that the template vector employed for HDR repair unwantedly and commonly inserts into the genome. We observe this deviation from the desired edit at multiple genomic locations, with different HDR vectors and with multiple genome editing designs. To avoid these events, we have generated a novel HDR template vector that enables animals with these undesired insertions to be identified and excluded. Our results suggest that HDR based genome edited animals must be carefully screened for unwanted vector template genomic integration in order to avoid misleading interpretations of genome editing outcomes.

Genome editing | CRISPR/Cas9 | *Drosophila*
Correspondence: *brian.mccabe@epfl.ch*
***These authors contributed equally to this study.**

Introduction

Genome engineering in model organisms and cell lines is a rapidly evolving field with continuously developing methods, techniques, and approaches (1, 2). The field was revolutionized by the employment of the clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 system, derived from prokaryotic immunity, which enables the precise targeting and subsequent modification of genes (3). The system forms the basis to create many types of genetic modifications, for example, insertions, deletions, or substitutions in genes of interest, with improved ease, speed, and accuracy (4). CRISPR/Cas9 based genome editing relies upon guide RNAs (gRNAs) to direct the endonuclease Cas9 protein to a target DNA sequence, leading to the introduction of double-stranded DNA breaks (DSBs) (5). The cell can then repair these DSBs by either nonhomologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ ligates nonhomologous DNA ends, which if occurring erroneously, can induce either insertions or deletions (indels) (6). HDR on the other hand, uses a DNA template from a homologous DNA strand to repair DSBs (7). Providing an exogenous 'template' sequence for HDR, commonly delivered from a plasmid vector, enables exploitation of the HDR process to create specific genome insertions, deletions or other nucleotide modifications, when these changes are introduced into the exogenous HDR template. The adaptability of the approach has enhanced the popularity of CRISPR/Cas9 as a technique for germline genome editing in model organisms (8) including *Drosophila melanogaster* (9). Germline genome editing in *Drosophila* employing HDR is most frequently carried out by germline injection of a HDR template plasmid, which contains homology arms flanking the desired edited sequence, together with plasmids expressing one or more gRNAs designed to induce DSBs in the targeted locus. To induce recombination, Cas9 is introduced in the germline either via transient expression from a plasmid or using transgenic animals stably expressing Cas9 (Figure 1) (10). Successfully edited animals can then be identified by molecular screening (e.g. PCR), phenotypic alterations or selectable markers (10).

While there are many advantages to the use of CRISPR/Cas9 based genome editing compared to other available methods, several studies have also revealed drawbacks of the technique. In mouse embryonic stem cells, hematopoietic progenitors and human differentiated cell lines, the DSBs created by guide RNAs and Cas9 have the potential to induce large deletions or rearrangements of the genomes (11). Similar undesired large deletions and insertions have been observed during mouse zygote genome editing (12). Here, we describe an undesired and highly frequent mutagenic insertion of the HDR donor plasmid template vector into the genome. Analogous gene editing problems have been previously identified (13–17), however, the frequency of these undesired events was not fully apparent. Here, we observe that HDR vector template insertion occurs to an alarmingly high degree with multiple template vectors. Our results caution investigators carrying out HDR directed genome editing in *Drosophila* to be on alert for erroneous template genomic insertion and mutagenesis, which may obscure and confound analysis of genome edited animals. To combat this problem, we have designed and tested a novel and versatile HDR template vector that allows the identification

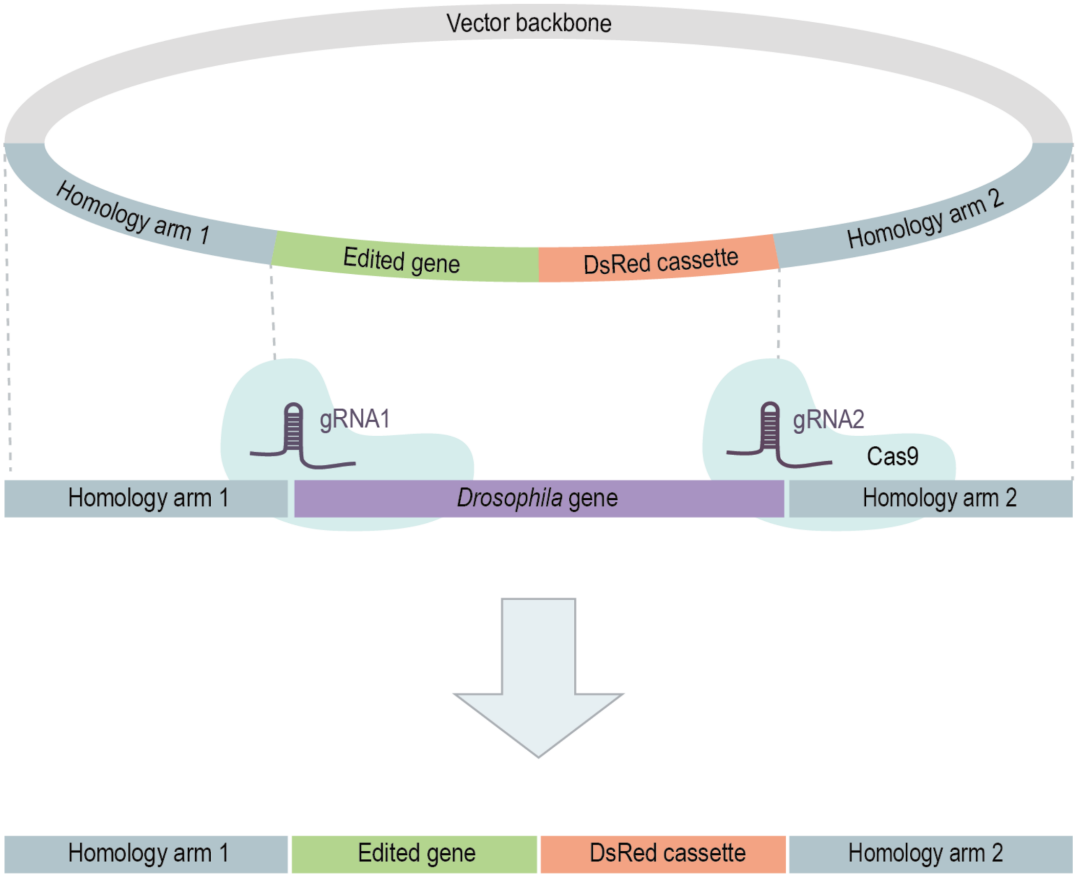

Figure 1. Schematic of the use of CRISPR/Cas9 HDR strategy. *Drosophila* embryos, expressing germline Cas9, are injected with plasmids expressing two gRNAs in addition to a plasmid containing the edited target sequence flanked by ∼1 kb long homology arms which act as a template for HDR. An eye expression DsRed selectable cassette is also included to facilitate identification of edited animals. After the cleavage of the targeted locus by gRNA/Cas9, homology to the exogenous template sequence enables HDR, resulting in the conversion of the genomic sequence to match the template plasmid sequence.

and exclusion of these undesired mutagenic insertional events, in addition to enabling 'all-in-one' template plus gRNA sequences single vector HDR genome editing.

Results

Undesired insertional mutagenesis induced during the generation of genome edited humanised *Drosophila*. The TAR DNA-binding protein 43 (*hTDP-43*) is associated with both sporadic and familial Amyotrophic lateral sclerosis (ALS) (18). With the goal to generate high fidelity *Drosophila* models of *hTDP-43* ALS, we sought to create Human Gene Replacement (HGR) or 'humanised' genome edited *Drosophila* lines, where we replaced the *Drosophila* TAR DNA-binding protein-43 homolog (*TBPH*) (19) gene with a human *hTDP-43* encoding cDNA (*hTDP-43*) (20). We began by using CRISPR Optimal Target Finder software (21), to design two gRNAs outside of the *TBPH* 3' and 5' UTR regions to induce Cas9 dependent DSBs. Both of the gRNA sequences we designed were synthesised and inserted into pCFD3 to facilitate expression in germline cells by transgene injection. To enable the replacement of the *Drosophila TBPH* open reading frame (ORF) with a human *hTDP-43* ORF while maintaining *Drosophila* 5' and 3' regulatory regions, we generated a HDR template plasmid containing ∼1 kb of 5' and 3' sequences homologous to the genomic region surrounding *Drosophila TBPH*, together with an artificial replacement exon encoding *hTDP-43*, and a selection cassette to allow identification of successfully edited animals by eye specific DsRed expression (Figure 2A) (22, 23). These sequences were cloned into the plasmid vector pCR4 (24). Both the gRNA encoding plasmids and the HDR replacement cassette were injected into *Drosophila* strains expressing Cas9 in the germline (25). Animals where genome editing had occurred were then identified by eye expression of DsRed. Other than the aim of *Drosophila* to human gene replacement, the design of both plasmids encoding the targeting gRNAs and containing the HDR template followed conventional practices and widely utilised approaches (9, 10).

After identification of genome edited animals, individual lines were produced, and homozygous animals from each

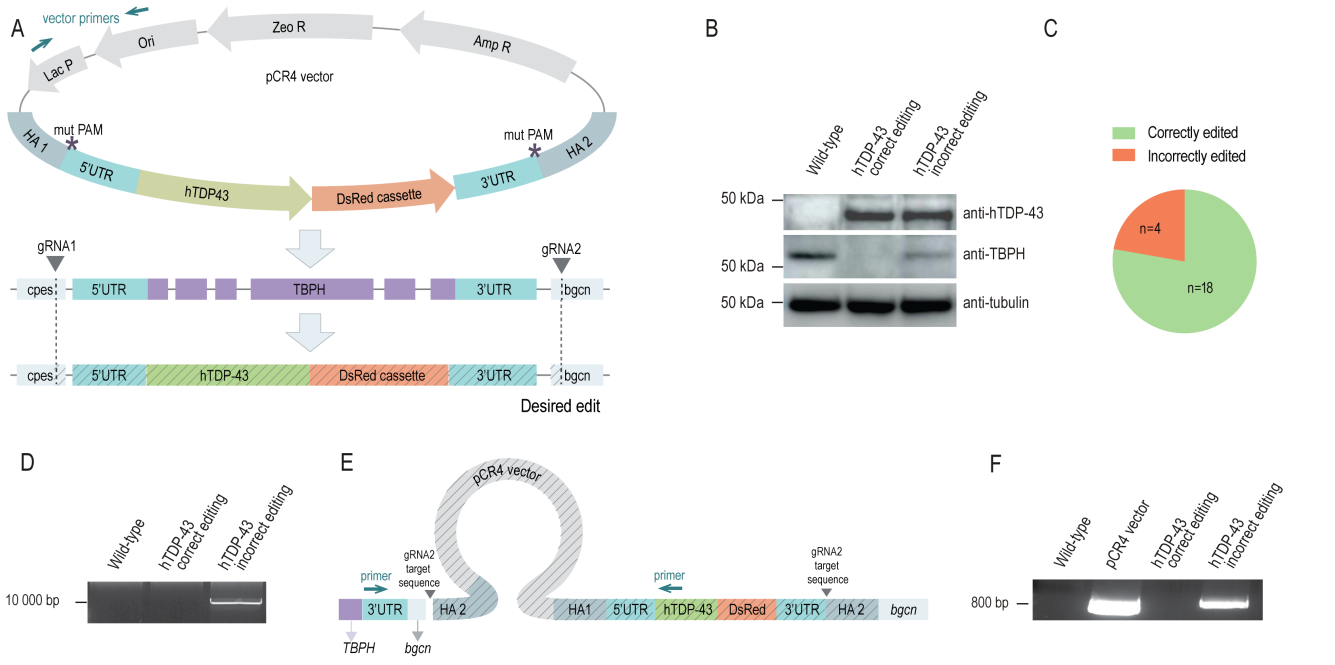

Figure 2. Undesired genomic insertion of HDR template vector sequence during gene replacement HDR. (A) Illustration of the CRISPR/Cas9 HDR replacement of the endogenous *Drosophila TBPH* gene by *hTDP-43*. A HDR template cloned into a circular pCR4 vector includes two ∼1 kb long homology arms on each side of *hTDP-43*, mutated gRNA PAM sequences [*] and a selectable DsRed cassette. Two gRNAs targeting sequences in the *TBPH* neighbouring genes, *bgcn* and *cpes*, enable Cas9 induced DSBs. The desired HDR editing result is a complete substitution of the *Drosophila TBPH* gene by *hTDP-43* plus insertion of the DsRed cassette. (B) Western blot showing the expression of hTDP-43, *Drosophila* TBPH and tubulin loading control in protein extracts derived from *Drosophila* controls, correctly edited humanised *hTDP-43 Drosophila*, and a humanised *hTDP-43 Drosophila* line with incorrect editing due to template genomic insertion. (C) Chart illustrating the fraction of humanised *hTDP-43 Drosophila* lines that resulted in an incorrect insertion of the template vector backbone (n=4), compared to the fraction of lines with *TBPH* correctly replaced by *hTDP-43* (n=18). (D) Image of an agarose gel from a genomic DNA PCR using a forward primer in the *TBPH* 3' end and a reverse primer in the *hTDP-43* 5' end, indicated in panel (E), which produces a product of ∼10 kb from genomic DNA derived from incorrectly edited *hTDP-43 Drosophila*, which is not present in genomic DNA from *Drosophila* controls or a correctly edited replacement *hTDP-43* line. (E) Schematic of the integrated vector sequence in incorrectly edited *hTDP-43* animals. The entire HDR template vector sequence is inserted into the genome and flanked by sequences matching homology arm 2. Primers used to detect the presence of the undesired vector backbone in the genome are represented by blue arrows. (F) Image of an agarose gel from genomic DNA amplification using primers specific for the template pCR4 vector. Vector primer positions are indicated in panel (A). Vector plasmid sequences can be detected in vector DNA and genomic DNA derived from incorrectly edited *hTDP-43* lines, but not in genomic DNA derived from correctly edited *hTDP-43 Drosophila* or controls. Diagrams are not to scale.

line were analysed for expression of the *Drosophila TBPH* and *hTDP-43* proteins. Our expectation was that animals would express only *hTDP-43* but not *Drosophila TBPH* if the editing event was successful. Indeed, using western blotting, we did identify 18 of 22 lines where only *hTDP-43* was expressed and *Drosophila TBPH* was undetectable. We subsequently confirmed in these lines using genomic DNA sequencing that the expected editing event had occurred.

However, we found that 18% (4 out of 22) of the lines we generated expressed both *Drosophila TBPH* and human *hTDP-43* (Figure 2B,C). To understand how this unwanted outcome could occur we investigated these lines further. We began by designing a forward primer in the *Drosophila TBPH* gene and a reverse primer for the *hTDP-43* gene in an attempt to amplify the region between these genes. Using genomic DNA generated from *Drosophila* lines expressing both *hTDP-43* and *TBPH* as a template, we amplified a ∼10 kb product that was not present in unedited animals or in animals where the editing event had occurred correctly (Figure 2D). We then sequenced this product and discovered that, in addition to including genomic sequences flanking *Drosophila TBPH*, it also included the entire pCR4 plasmid vector sequence, indicating that the entire pCR4 HDR editing template plasmid vector had inserted into the genome of these animals (Figure 2E). This insertion occurred proximal to *Drosophila TBPH* in the region between *TBPH* and the nearby gene *bgcn*. We examined the precise sequence where both the 5' and 3'gRNAs were designed to target in this sequence. We found that the donor vector had inserted into the genome at the location where a DSB was designed to have been induced by one of the gRNAs. At this location, on both sides of the insertion, the PAM sequence of the target region was

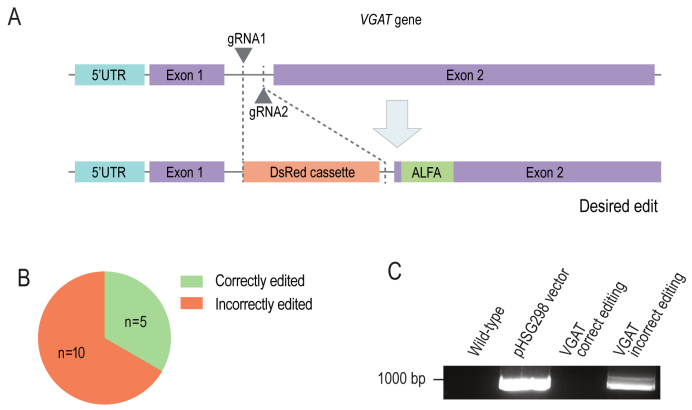

Figure 3. Undesired template vector integration in genome editing to insert an epitope tag. (A) Strategy for CRISPR/Cas9 plus HDR insertion of an ALFA tag in frame in exon 2 in the *VGAT* gene. Two gRNAs enable two DSBs between exons 1 and 2 of the *VGAT*. The ALFA tag and a removable DsRed cassette, flanked by ∼1 kb long homology arms, in a pHSG298 vector acts as a template for HDR. Correct editing results in a DsRed cassette in intron 1 and an in-frame ALFA tag sequence within exon 2. (B) Chart of the ratio of correctly inserted tagged *VGAT Drosophila* lines, compared to those with inserted template vector sequence (C) Image of an agarose gel of products amplified using PCR using primers specific for the vector pHSG298. pHSG298 sequences can be amplified from pHSG298 and genomic DNA derived from unwanted editing of *VGAT Drosophila* lines, but not correctly edited *VGAT Drosophila* lines or controls. Diagrams are not to scale.

modified to match the edited PAM sequence of the template donor plasmid consistent with this being the insertion point of the vector sequence. From this we conclude that the entire template vector has inserted into the genome at the site where a DSB was designed to be induced by one of the gRNAs. To determine if the HDR donor vector sequence was present in all incorrectly edited animals, we designed primers to amplify the region between the lac promoter and pUC ori sequence of pCR4. Using these primers, we found that 100% (4 out of 4) of incorrectly edited animals, i.e. where *Drosophila TBPH* and *hTDP-43* were both expressed, had pCR4 sequences inserted into their genome. No product could be amplified from the genomic DNA derived from background control or correctly edited animals (Figure 2F). Thus, in these anomalously edited animals, we observed an undesired incorporation of the entire HDR donor vector into the genome. We also observed that the undesired template vector insertion could be inherited stably in animals across multiple generations consistent with germline genome insertion.

Unwanted HDR vector genomic incorporation during insertion of sequences into an endogenous gene. We wondered if the high rate of insertion of vector sequences into the genome during *TBPH* editing was associated with the particular gRNAs we had utilised for this HGR editing procedure, or due to the nature of the large editing event that we had attempted to achieve. To investigate this further, we examined a different type of editing event, i.e. in a different target gene and employing different gRNAs. The Vesicular GABA transporter (*VGAT*) (26) is required for the release of the neurotransmitter GABA. We wished to insert an ALFA epitope tag (27) in frame into the *VGAT* ORF to enable analysis of protein expression. Again, we designed two gRNAs using CRISPR Optimal Target Finder and an HDR donor plasmid consisting of homology arms, *VGAT* sequence that included the epitope tag and an eye DsRed selection cassette, but employed a different cloning vector, pHSG298 (Figure 3A) guided by a prior studies (28). Fifteen genome edited animals were recovered with eye DsRed expression and the presence of the ALFA tag was assessed by PCR genotyping. Using primers specific for the ALFA tag and *VGAT* flanking regions outside of the region of HDR homology arms, we confirmed the presence of an ALFA tag insertion in all lines. We then tested how many of these lines had undesired insertion of pHSG298 sequences. We found that 10 lines (66.6%) had incorporation of pHSG298 sequences into the genome, with only 5 edited (33.3%) lines showing no evidence of a donor vector insertional event (Figure 3B,C). Importantly, we examined the lines with undesired vector incorporation using immunohistochemistry and western blotting to confirm correct ALFA tag insertion into *VGAT*. We confirmed that indeed several lines did have ALFA tag insertion into the *VGAT* gene (data not shown). Therefore, unlike TPBH gene replacement described above, correct *VGAT* genome editing could occur in tandem with the unwanted insertion of the HDR template vector sequence into the genome. Our results also indicated that the choice of HDR donor vector does not seem to be critical, as pHSG298 sequences were inserted similarly to pCR4. Thus, in an independent HDR based genome edit of an endogenous *Drosophila* gene, we also observed digressive template vector genomic incorporation.

HDR template vector mutagenic insertion editing a viable gene. Both *TBPH* and *VGAT* are recessive lethal when mutated (29, 30). We wondered, even though selection for editing events is carried out in heterozygotes, if this could in some way bias towards undesired HDR template vector genome incorporation. Also, both of the genome edits described above included insertion of exogenous sequences (human *hTDP-43*, ALFA epitope tag) into the genome. We therefore examined another HDR editing event, where our aim was to mutate nucleotides within the Glutamate Receptor IIA (*GluRIIA*) gene, which is homozygous viable even in protein null mutants (31). Similar to the above genome editing strategies, we designed two gRNAs to target exon 12 of *GluRIIA* and an HDR vector (in pCR4) with homology arms designed to replace

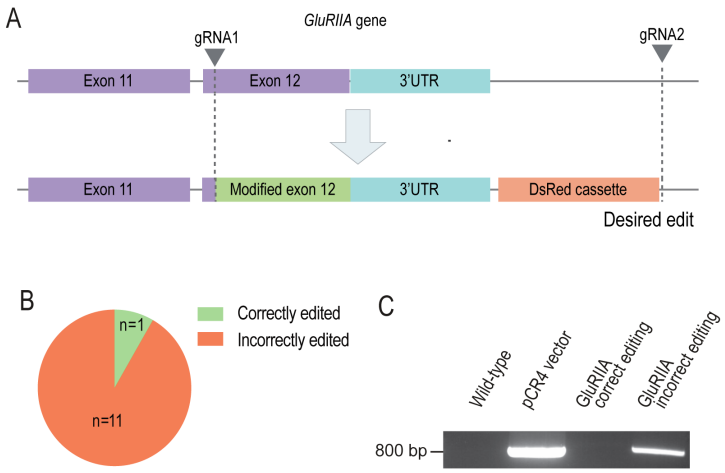

Figure 4. Undesired HDR vector genomic integration in genome editing to alter single nucleotides. Illustration of the CRISPR/Cas9 plus HDR editing strategy to edit single nucleotides in exon 12 of the *GluRIIA* gene by employing two gRNAs to enable DSBs. A pCR4 plasmid containing the mutated *GluRIIA* exon 12 and a removable DsRed cassette, flanked by ∼1 kb long homology arms was provided a template for HDR. Correct editing results in a modified version of exon 12 labelled by a selectable DsRed cassette. (B) Chart illustrating the number of lines with correct exon replacement (n=1) compared to the number of lines with incorrect template vector backbone insertion (n=11). (C) Agarose gel image of genomic DNA PCR amplification using primers specific for pCR4. The vector backbone sequence was identified in incorrectly edited *GluRIIA Drosophila* lines, but not in the correctly edited *GluRIIA Drosophila* line or controls. Diagrams are not to scale.

two amino acids in *GluRIIA* with modified variants, again using eye DsRed expression to select editing events (Figure 4A). Twelve DsRed expressing *Drosophila* lines were recovered and were analysed by PCR genomic amplification and sequencing. Of these 12 lines, only one had the correct edit, as confirmed by DNA sequencing (Figure 4B). From the other 11 lines, DNA sequencing showed two peaks in the sequencing chromatogram, revealing the presence of the unmutated *GluRIIA* nucleotides in addition to the desired modified nucleotides. This suggested that both variants of this exon sequence were present in the genome of these animals. To confirm that this was due to insertion of the HDR vector into the genome, we carried out a PCR for pCR4 vector sequences from genomic DNA derived from these incorrectly edited animals (Figure 4C). All 11 incorrectly edited lines were confirmed to have pCR4 sequences inserted into their genome. Thus, again, with a third set of independent gRNAs targeting another independent gene, we observed a high frequency of anomalous template donor vector insertion into the genome as a problematic consequence of HDR based genome editing.

Undesired insertion of HDR template vector can be excluded through use of a selectable marker. Selection markers are widely used in yeast, mice and *Drosophila* to detect exogenous sequence insertional events (13–15, 32, 33). As we wished to continue to use a DsRed piggyBAC selection cassette (22) to identify successful editing events (and subsequently 'scarlessly' remove it after successful edit identification), we sought to incorporate an additional selection marker into the HDR template backbone to identify and exclude vector insertional mutagenic events from subsequent analysis. Mini-white (34) is a commonly used selectable marker in *Drosophila*, and in addition to eyes absent selection, has previously been evaluated as a template vector negative selection marker (35). However mini-white selection proved unreliable for detection of vector genomic insertion in our hands (data not shown). We therefore decided to design a vector using a completely exogenous gene for negative selection with a bright selection marker that would be easy and quick to score. We designed a HDR template cloning vector to produce expression of ZsGreen1 (36), a very bright green fluorescent protein, under the eye P3 promoter (37) only if the vector backbone becomes incorporated into the genome. We dubbed this vector pVID - Vector Insertion Detector (Figure 5A).

To evaluate the effectiveness of pVID, we performed genome editing of *Synapsin* (38), which encodes a phosphoprotein involved in vesicle release, designed to insert ALFA epitope tags within the *Synapsin* open reading frame. Similar to the genome editing procedure described above, we designed two gRNAs using CRISPR Optimal Target Finder and an HDR donor plasmid consisting of homology arms, a *Synapsin* sequence that included the epitope tags and an eye DsRed selection cassette, but employed pVID as the cloning vector (Figure 5B). 16 DsRed expressing *Drosophila* lines were recovered. Of these, 8 (50%) also had ZsGreen expression. We carried out PCR for pVID vector sequences from genomic DNA derived from ZsGreen expressing lines and found all (8 of 8) had template vector insertion. In contrast, PCR for pVID vector sequences from genomic DNA derived from DsRed expressing lines which did not have ZsGreen expression, showed no evidence of vector template sequences (8 of 8), plus all had the expected *Synapsin* genome edit. We additionally utilised pVID for second CRISPR/Cas9 HDR genome editing procedure to insert sequences into a different gene, GluRIIC, which encodes an obligate Glutamate receptor (39). The GluRIIC edit included a HDR template with the desired insertion and an eye DsRed selection cassette. As was the case for *Synapsin* editing, we recovered lines with both DsRed and ZsGreen eye expression (5 of 6) and a line with DsRed eye expression but no ZsGreen eye expression (1 of 6). We carried out PCR for pVID vector sequences from genomic DNA derived from ZsGreen expressing

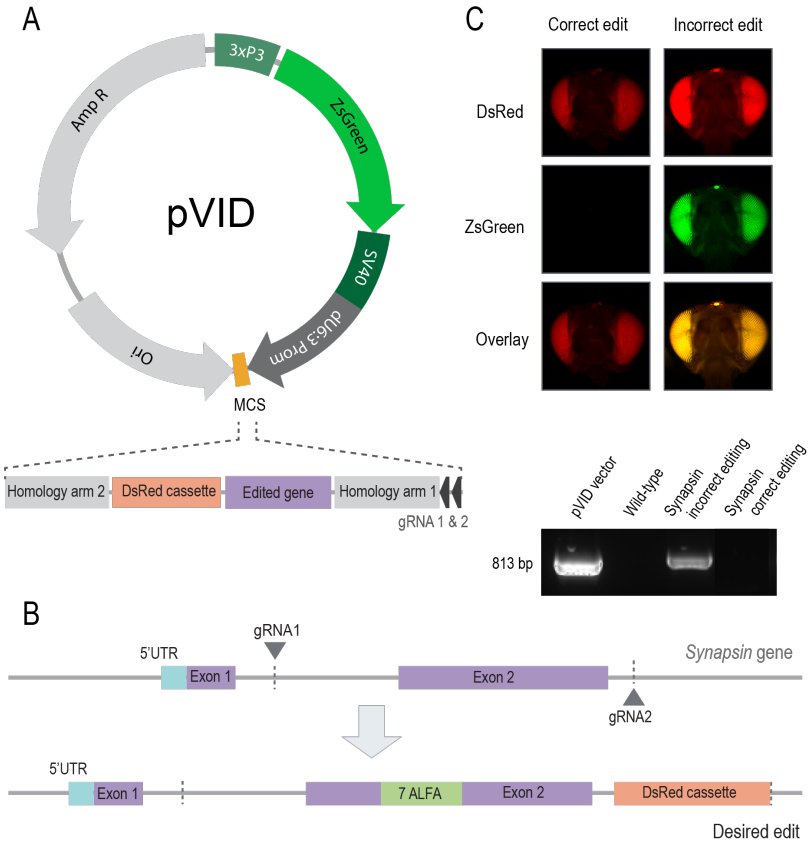

Figure 5. The pVID HDR template vector enables the identification of *Drosophila* genome edited lines with undesired vector integration events using ZsGreen. A) Graphical map of pVID vector (to scale). It includes an Ampicillin resistance, a *Drosophila* U6:3 promotor for gRNA expression (dark grey), a multiple cloning site (MCS) for the insertion of the HDR template sequence (orange) and a ZsGreen selection cassette (green) with its 3xP3 eye promoter, ZsGreen fluorescent protein coding sequence and SV40 terminator. In our editing strategy, the HDR template sequence contains the desired edit and a DsRed cassette selection flanked by two homology arms, plus two gRNAs under the control of dU6:3 promoter (dark grey). B) Strategy for CRISPR/Cas9 plus HDR insertion of seven ALFA tags in frame in exon 2 in the *Synapsin* gene. Two gRNAs enable two DSBs between exons 1 and 2 and after exon 2. Correct editing results in a DsRed cassette in intron 2 and an in-frame 7ALFA tag sequence within exon 2.(C) Image of eye scoring for correct and incorrect edit. DsRed positive eye expression indicates the presence of the desired genomic edit, DsRed and ZsGreen eye expression indicate that edit is accompanied by an undesired vector backbone insertion. (D) Agarose gel image of pVID vector backbone PCR amplification from edited *Synapsin Drosophila* lines and controls, showing undesired vector insertion within the genome of ZsGreen expressing animals. Diagrams are not to scale.

lines and found all (5 of 5) had HDR template vector insertion. In contrast, the line which only expressed DsRed had no evidence of HDR template vector insertion. We conclude that use of pVID as the backbone vector for HDR templates utilised for genome editing allows the successful identification and exclusion of undesired mutagenic HDR vector insertion events.

Discussion

CRISPR/Cas9 based genome editing has been widely deployed to introduce, remove or modify genetic sequences in many model organisms including *Drosophila* (10, 23, 40). Here we describe the highly frequent undesired insertion of the donor template vector into the genome in experiments designed to alter genes based upon homology directed repair. We observed this unwanted insertional mutagenesis in five fully independent HDR procedures utilising *Drosophila* germline genome editing. Looking across the totality of the genome edits described here, an average of 54% (n=71) of the edited lines we generated had undesired vector insertion, consistent with other reports (17). The procedures we employed to design our HDR edits were not unusual. gRNA sequences were designed using CRISPR Optimal Target Finder based on conventional on-target and off target selection criteria (21) (Table 1). The germline Cas9 expressing animals we employed are also widely used (25). Our HDR repair templates utilised conventional design with ∼1 kb flanking homology arms together with the sequence to be inserted or edited including mutated PAM sequences (9). Slightly less common, yet also widely employed, was our use of an eye expression DsRed selection cassette to identify editing events, thanks to the ability to 'scarlessly' remove it after successful transgene or genome edit identification (22). Indeed, the presence of DsRed allowed us to unambiguously identify events where the genome was edited. HDR donor vectors that do not include a visible maker also integrate into the genome at high frequencies but are much harder to detect (17). The precise sequence of the HDR vector also does not seem to influence genome integration as we observed insertion of both pCR4 (24), pHSG298 (41) and, as described here, pVID.

Integration of the HDR template is a major issue for subsequent analysis of edited animals, potentially giving rise to misleading results. For example, if the genome edit has not occurred as designed but is falsely indicated as correct, by genomic PCR from the integrated vector, results may be misinterpreted. Alternatively, mutagenic integration of the vector sequence may cause phenotypes not attributable to the designed edit. Unexpected insertions have previously been observed during CRISPR/Cas9 based mouse genome editing, at a rate of 10% when targeting individual genomic sites with more than one gRNA, compared to 4% when targeting individual genomic sites with a single gRNA (12).

In our experiments, we speculate that one of the gRNAs we designed to target a gene has a lower efficacy leading recombination to only occur at one gRNA site instead of two, causing a single homologous recombination event and insertion of the vector template into the locus. Indeed, it has been reported that gRNAs can have differing levels of efficiency which can affect editing throughput (42, 43). It is possible, but untested by us, that HDR strategies based upon the use of a single guide RNA might reduce undesired HDR donor vector insertion (17, 44) though perhaps at the cost of efficiency (45).

It is also possible that other factors, besides the efficiency of the gRNAs, enable undesired vector backbone insertion. One possibility is that the HDR vector template becomes linearised at some frequency which could facilitate insertion into the genome. Linearized templates are becoming increasingly employed in the field of CRISPR/Cas9 in several models, due to increased efficiency of editing (44, 46, 47). However, it is noteworthy that during HDR germline editing of mice, insertion of linear HDR template sequences has also been observed, not just in one copy but in several head-to-tail copies (48). These insertions were shown to be produced by both HDR and NHEJ, or in some cases, a combination of both mechanisms. Similar mechanisms could be compatible with the insertion events that we observe. Therefore, both with linear and circular templates, it is important to remain vigilant by including controls for undesired vector template insertions. Consistent in all of our experiments is also the fact that we express Cas9 in the *Drosophila* germ cells under the nos promoter to enable HDR to occur. It is possible that other Cas9 sources, such as Cas9 driven by a different promoter, or at a lower or higher expression level, could influence the insertion of the backbone vector though very high levels of Cas9 can be toxic (49).

In order to improve efficiency of HDR-mediated genome editing in our laboratory, we now routinely use two selection cassettes, a positive selection (subsequently removable) dsRed cassette to indicate an editing event has occurred and a negative selection ZsGreen cassette to eliminate events where the HDR template vector has inserted into the genome. We would suggest that others using similar approaches might consider the same, which unfortunately, since HDR template vector insertion is as we describe common, requires the generation of many additional lines to select only those that have the desired edit. Ultimately however, it would be desirable to understand more about the factors that lead to undesired HDR vector genomic integration, in order to design strategies to avoid this unwanted, labor adding and detrimental outcome of CRISPR/Cas9 based HDR dependent genome editing.

Limitations of the Study

We note that while we observed the insertion of the template vector using three different vector template backbones, there were still similarities between the five different *Drosophila* lines we analysed in this paper. For example, in all lines, two gRNAs and a removable DsRed selection cassette were employed. In future experiments it would be beneficial to investigate the possible presence of template vector in lines which have been produced by a different number of gRNAs using a different selection marker. We also tested for undesired insertional mutagenesis only during *Drosophila* melanogaster germline genome editing and not in other model organisms or cell lines where, given similar procedures, this undesired outcome could also potentially occur. In addition, we used exclusively the nos promoter to highly express Cas9 in *Drosophila* germ cells, and it is possible that the Cas9 concentration could affect vector template integration.

Acknowledgements

We would like to thank Dr. Steven Stowers for assistance with vectors and the design of the *VGAT* targeting strategy plus for critical reading of the manuscript. We also are grateful to other members of the McCabe laboratory for their criticisms. EK was funded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 721802. MR was funded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 945363. This work was supported by the Swiss National Science Foundation grant numbers: CRSII3_154455, 31003A_179587 and 320030-232324 to B.M.

Author Contributions

Conceptualization, B.M.; Investigation, E.K., E.R, M.R, J.A.; Writing – Original Draft, E.K. and B.M.; Writing – Review & Editing, E.K., B.M., E.R., M.R.; Funding acquisition, B.S. and B.M.

Declaration of Interests

The authors declare no competing interests.

Bibliography

1. Marcin G. Fraczek, Samina Naseeb, and Daniela Delneri. History of genome editing in yeast. *Yeast (Chichester, England)*, 35(5):361–368, May 2018.
2. Dagmara Korona, Stefan A. Koestler, and Steven Russell. Engineering the Drosophila Genome for Developmental Biology. *Journal of Developmental Biology*, 5(4):16, December 2017. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
3. Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, 337(6096):816–821, August 2012.
4. Yue Yang, Jin Xu, Shuyu Ge, and Liqin Lai. CRISPR/Cas: Advances, Limitations, and Applications for Precision Cancer Research. *Frontiers in Medicine*, 8, 2021.
5. Han Yang, Shuling Ren, Siyuan Yu, Haifeng Pan, Tingdong Li, Shengxiang Ge, Jun Zhang, and Ningshao Xia. Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks. *International Journal of Molecular Sciences*, 21(18):6461, September 2020.
6. Howard H. Y. Chang, Nicholas R. Pannunzio, Noritaka Adachi, and Michael R. Lieber. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology*, 18(8):495–506, August 2017. Number: 8 Publisher: Nature Publishing Group.
7. Claire Wyman and Roland Kanaar. DNA double-strand break repair: all's well that ends well. *Annual Review of Genetics*, 40:363–383, 2006.
8. Dongyuan Ma and Feng Liu. Genome Editing and Its Applications in Model Organisms. *Genomics, Proteomics & Bioinformatics*, 13(6):336–344, December 2015.
9. Fillip Port and Simon L. Bullock. Creating Heritable Mutations in Drosophila with CRISPR-Cas9. In Christian Dahmann, editor, *Drosophila: Methods and Protocols*, Methods in Molecular Biology, pages 145–160. Springer, New York, NY, 2016.
10. Scott J. Gratz, C. Dustin Rubinstein, Melissa M. Harrison, Jill Wildonger, and Kate M. O'Connor-Giles. CRISPR-Cas9 genome editing in Drosophila. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]*, 111:31.2.1–31.2.20, July 2015.
11. Michael Kosicki, Kärt Tomberg, and Allan Bradley. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nature Biotechnology*, 36(8):765–771, September 2018. Number: 8 Publisher: Nature Publishing Group.
12. Ha Youn Shin, Chaochen Wang, Hye Kyung Lee, Kyung Hyun Yoo, Xianke Zeng, Tyler Kuhns, Chul Min Yang, Teresa Mohr, Chengyu Liu, and Lothar Hennighausen. CRISPR/Cas9 targeting events cause complex deletions and insertions at 17 sites in the mouse genome. *Nature Communications*, 8(1):15464, May 2017. Number: 1 Publisher: Nature Publishing Group.
13. Lino Tessarollo, Mary Ellen Palko, Keiko Akagi, and Vincenzo Coppola. Gene Targeting in Mouse Embryonic Stem Cells. *Methods in molecular biology (Clifton, N.J.)*, 530:141–164, 2009.
14. Raphaël Dutoit, Evelyne Dubois, and Eric Jacobs. Selection systems based on dominant-negative transcription factors for precise genetic engineering. *Nucleic Acids Research*, 38(19):e183, October 2010.
15. Juan Huang, Wenke Zhou, Annie M Watson, Yuh-Nung Jan, and Yang Hong. Efficient Ends-Out Gene Targeting In Drosophila. *Genetics*, 180(1):703–707, September 2008.
16. Zhongsheng Yu, Hanqing Chen, Jiyong Liu, Hongtao Zhang, Yan Yan, Nannan Zhu, Yawen Guo, Bo Yang, Yan Chang, Fei Dai, Xuehong Liang, Yixu Chen, Yan Shen, Wu-Min Deng, Jianming Chen, Bo Zhang, Changqing Li, and Renjie Jiao. Various applications of TALEN- and CRISPR/Cas9-mediated homologous recombination to modify the *Drosophila* genome. *Biology Open*, 3(4):271–280, April 2014.
17. Daniel Tianfang Ge, Cindy Tipping, Michael H Brodsky, and Phillip D Zamore. Rapid Screening for CRISPR-Directed Editing of the *Drosophila* Genome Using *white* Coconversion. *G3 Genes|Genomes|Genetics*, 6(10):3197–3206, October 2016.
18. Ian R.A. Mackenzie and Rosa Rademakers. The role of TDP-43 in amyotrophic lateral sclerosis and frontotemporal dementia. *Current opinion in neurology*, 21(6):693–700, December 2008.
19. Fabian Feiguin, Vinay K. Godena, Giulia Romano, Andrea D'Ambrogio, Raffaella Klima, and Francisco E. Baralle. Depletion of TDP-43 affects Drosophila motoneurons terminal synapsis and locomotive behavior. *FEBS letters*, 583(10):1586–1592, May 2009.
20. Jer-Cherng Chang and David B. Morton. Drosophila lines with mutant and wild type human TDP-43 replacing the endogenous gene reveals phosphorylation and ubiquitination in mutant lines in the absence of viability or lifespan defects. *PLOS ONE*, 12(7):e0180828, July 2017. Publisher: Public Library of Science.
21. Scott J. Gratz, Fiona P. Ukken, C. Dustin Rubinstein, Gene Thiede, Laura K. Donohue, Alexander M. Cummings, and Kate M. O'Connor-Giles. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics*, 196(4):961–971, April 2014.
22. Ethan Bier, Melissa M Harrison, Kate M O'Connor-Giles, and Jill Wildonger. Advances in Engineering the Fly Genome with the CRISPR-Cas System. *Genetics*, 208(1):1–18, January 2018.
23. Fillip Port, Hui-Min Chen, Tzumin Lee, and Simon L. Bullock. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proceedings of the National Academy of Sciences*, 111(29):E2967–E2976, July 2014. Publisher: Proceedings of the National Academy of Sciences.
24. S. Shuman. Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *The Journal of Biological Chemistry*, 269(51):32678–32684, December 1994.
25. Xingjie Ren, Jin Sun, Benjamin E. Housden, Yanhui Hu, Charles Roesel, Shuailiang Lin, Lu-Ping Liu, Zhihao Yang, Decai Mao, Lingzhu Sun, Qujie Wu, Jun-Yuan Ji, Jianzhong Xi, Stephanie E. Mohr, Jiang Xu, Norbert Perrimon, and Jian-Quan Ni. Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. *Proceedings of the National Academy of Sciences*, 110(47):19012–19017, November 2013. Publisher: Proceedings of the National Academy of Sciences.
26. Lina Enell, Yasutaka Hamasaka, Agata Kolodziejczyk, and Dick R. Nässel. γ-Aminobutyric acid (GABA) signaling components in Drosophila: Immunocytochemical localization of GABAB receptors in relation to the GABAA receptor subunit RDL and a vesicular GABA transporter. *Journal of Comparative Neurology*, 505(1):18–31, 2007. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.21472.
27. Hansjörg Götzke, Markus Kilisch, Markel Martínez-Carranza, Shama Sograte-Idrissi, Abirami Rajavel, Thomas Schlichthaerle, Niklas Engels, Ralf Jungmann, Pål Stenmark, Felipe Opazo, and Steffen Frey. The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. *Nature Communications*, 10(1):4403, September 2019. Number: 1 Publisher: Nature Publishing Group.
28. Sarah J. Certel, Brian D. McCabe, and R. Steven Stowers. A conditional GABAergic synaptic vesicle marker for Drosophila. *Journal of Neuroscience Methods*, 372:109540, April 2022.
29. Danielle C. Diaper, Yoshitsugu Adachi, Ben Sutcliffe, Dickon M. Humphrey, Christopher J.H. Elliott, Alan Stepto, Zoe N. Ludlow, Lies Vanden Broeck, Patrick Callaerts, Bart Dermaut, Ammar Al-Chalabi, Christopher E. Shaw, Iain M. Robinson, and Frank Hirth. Loss and gain of Drosophila TDP-43 impair synaptic efficacy and motor control leading to age-related neurodegeneration by loss-of-function phenotypes. *Human Molecular Genetics*, 22(8):1539–1557, April 2013.
30. Hao Fei, Dawnis M. Chow, Audrey Chen, Rafael Romero-Calderón, Wei S. Ong, Larry C. Ackerson, Nigel T. Maidment, Julie H. Simpson, Mark A. Frye, and David E. Krantz. Mutation of the Drosophila vesicular GABA transporter disrupts visual figure detection. *The Journal of Experimental Biology*, 213(10):1717–1730, May 2010.
31. Sophie A. Petersen, Richard D. Fetter, Jasprina N. Noordermeer, Corey S. Goodman, and Aaron DiAntonio. Genetic Analysis of Glutamate Receptors in Drosophila Reveals a Retrograde Signal Regulating Presynaptic Transmitter Release. *Neuron*, 19(6):1237–1248, December 1997. Publisher: Elsevier.
32. Gabriel T. Koreman, Yineng Xu, Qinan Hu, Zijing Zhang, Sarah E. Allen, Mariana F. Wolfner, Bei Wang, and Chun Han. Upgraded CRISPR/Cas9 tools for tissue-specific mutagenesis in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 118(14):e2014255118, April 2021.
33. Sasidhar Karuparti, Ann T. Yeung, Bei Wang, Pedro F. Guicardi, and Chun Han. A toolkit for converting Gal4 into LexA and Flippase transgenes in Drosophila. *G3 (Bethesda, Md.)*, 13(3):jkad003, March 2023.
34. Vincenzo Pirrotta. Vectors for P-Mediated Transformation in *Drosophila*. In Raymond L. Rodriguez and David T. Denhardt, editors, *Vectors*, pages 437–456. Butterworth-Heinemann, January 1988.
35. Kevin G. Nyberg, Joseph Q. Nguyen, Yong-Jae Kwon, Shelby Blythe, Greg J. Beitel, and Richard Carthew. A pipeline for precise and efficient genome editing by sgRNA-Cas9 RNPs in *Drosophila*. *Fly*, 14(1-4):34–48, October 2020.
36. M. V. Matz, A. F. Fradkov, Y. A. Labas, A. P. Savitsky, A. G. Zaraisky, M. L. Markelov, and S. A. Lukyanov. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature Biotechnology*, 17(10):969–973, October 1999.
37. G. Sheng, E. Thouvenot, D. Schmucker, D. S. Wilson, and C. Desplan. Direct regulation of rhodopsin 1 by Pax-6/eyeless in Drosophila: evidence for a conserved function in photoreceptors. *Genes & Development*, 11(9):1122–1131, May 1997.
38. Bert R. E. Klagges, Gertrud Heimbeck, Tanja A. Godenschwege, Alois Hofbauer, Gert O. Pflugfelder, Rita Reifegerste, Dietmar Reisch, Michael Schaupp, Sigrid Buchner, and Erich Buchner. Invertebrate Synapsins: A Single Gene Codes for Several Isoforms in Drosophila. *The Journal of Neuroscience*, 16(10):3154–3165, May 1996.
39. Scott B. Marrus, Scott L. Portman, Marcus J. Allen, Kevin G. Moffat, and Aaron DiAntonio. Differential localization of glutamate receptor subunits at the Drosophila neuromuscular junction. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 24(6):1406–1415, February 2004.
40. Mazhar Adli. The CRISPR tool kit for genome editing and beyond. *Nature Communications*, 9(1):1911, May 2018. Number: 1 Publisher: Nature Publishing Group.
41. Sunao Takeshita, Masahiro Sato, Mari Toba, Wakako Masahashi, and Tamotsu Hashimoto-Gotoh. High-copy-number and low-copy-number plasmid vectors for lacZα-complementation and chloramphenicol- or kanamycin-resistance selection. *Gene*, 61(1):63–74, January 1987.
42. Tobias Bruegmann, Khira Deecke, and Matthias Fladung. Evaluating the Efficiency of gRNAs in CRISPR/Cas9 Mediated Genome Editing in Poplars. *International Journal of Molecular Sciences*, 20(15):3623, July 2019.
43. Summer B. Thyme, Laila Akhmetova, Tessa G. Montague, Eivind Valen, and Alexander F. Schier. Internal guide RNA interactions interfere with Cas9-mediated cleavage. *Nature Communications*, 7:11750, June 2016.
44. Oguz Kanca, Jonathan Zirin, Jorge Garcia-Marques, Shannon Marie Knight, Donghui Yang-Zhou, Gabriel Amador, Hyunglok Chung, Zhongyuan Zuo, Liwen Ma, Yuchun He, Wen-Wen Lin, Ying Fang, Ming Ge, Shinya Yamamoto, Karen L Schulze, Yanhui Hu, Allan C Spradling, Stephanie E Mohr, Norbert Perrimon, and Hugo J Bellen. An efficient CRISPR-based strategy to insert small and large fragments of DNA using short homology arms. *eLife*, 8:e51539, 2019.
45. Nicholas S. McCarty, Alicia E. Graham, Lucie Studená, and Rodrigo Ledesma-Amaro. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nature Communications*, 11(1):1281, March 2020. Number: 1 Publisher: Nature Publishing Group.

46. Naoaki Shinzawa, Tsubasa Nishi, Fumiya Hiyoshi, Daisuke Motooka, Masao Yuda, and Shiroh Iwanaga. Improvement of CRISPR/Cas9 system by transfecting Cas9-expressing Plasmodium berghei with linear donor template. *Communications Biology*, 3(1):1–13, August 2020. Publisher: Nature Publishing Group.
47. Fei Song and Knut Stieger. Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks. *Molecular Therapy. Nucleic Acids*, 7:53–60, June 2017.
48. Boris V. Skryabin, Delf-Magnus Kummerfeld, Leonid Gubar, Birte Seeger, Helena Kaiser, Anja Stegemann, Johannes Roth, Sven G. Meuth, Hermann Pavenstädt, Joanna Sherwood, Thomas Pap, Roland Wedlich-Söldner, Cord Sunderkötter, Yuri B. Schwartz, Juergen Brosius, and Timofey S. Rozhdestvensky. Pervasive head-to-tail insertions of DNA templates mask desired CRISPR-Cas9–mediated genome editing events. *Science Advances*, 6(7):eaax2941, February 2020. Publisher: American Association for the Advancement of Science.
49. Nhan Huynh, Jie Zeng, Wen Liu, and Kirst King-Jones. A *Drosophila* CRISPR/Cas9 Toolkit for Conditionally Manipulating Gene Expression in the Prothoracic Gland as a Test Case for Polytene Tissues. *G3 Genes|Genomes|Genetics*, 8(11):3593–3605, November 2018.
50. Fillip Port and Simon L. Bullock. Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. *Nature Methods*, 13(10):852–854, October 2016.
51. Oguz Kanca, Jonathan Zirin, Yanhui Hu, Burak Tepe, Debdeep Dutta, Wen-Wen Lin, Liwen Ma, Ming Ge, Zhongyuan Zuo, Lu-Ping Liu, Robert W Levis, Norbert Perrimon, and Hugo J Bellen. An expanded toolkit for Drosophila gene tagging using synthesized homology donor constructs for CRISPR-mediated homologous recombination. *eLife*, 11:e76077, June 2022. Publisher: eLife Sciences Publications, Ltd.
52. Abhimanyu Sarkar, Asela Atapattu, Esther J. Belikoff, Jörg C. Heinrich, Xuelei Li, Carsten Horn, Ernst A. Wimmer, and Maxwell J. Scott. Insulated piggyBac vectors for insect transgenesis. *BMC Biotechnology*, 6(1):27, June 2006.
53. Yasuyuki Nakamura, Jun Ishii, and Akihiko Kondo. Bright fluorescence monitoring system utilizing Zoanthus sp. green fluorescent protein (ZsGreen) for human G-protein-coupled receptor signaling in microbial yeast cells. *PloS One*, 8(12):e82237, 2013.
54. Pascal Lehwark and Stephan Greiner. GB2sequin - A file converter preparing custom GenBank files for database submission. *Genomics*, 111(4):759–761, July 2019.

Materials and Methods

Plasmid construction. For all plasmids, homology arms, the desired edits, and the selection cassettes were PCR-amplified using a high-fidelity DNA polymerase (Platinum SuperFi II, Invitrogen) and assembled via HiFi DNA Assembly (Invitrogen), unless stated otherwise. All gRNAs were chosen using CRISPR Optimal Target Finder (21).

hTDP-43 knock-in HDR construct in pCR4. The full assembly was inserted by Topo cloning into pCR4. The two gRNAs cloned each in a pCDF3 vector (gRNA sequences in Table 1) targeted neighbouring genes cpes and bgcn, while making sure to not change the amino acid sequence of either gene. The inserted sequence included both *TBPH* 5' and 3'UTRs but with the *TBPH* coding sequence replaced by a *Drosophila* codon optimised version of the human *hTDP-43* cDNA. The whole edit was surrounded by ∼1 kb homology arms.

ALFA-VGAT tagged HDR construct in pHSG298. The cloning strategy was based on previous vGat genome editing (28). The full assembly was inserted by restriction digestion - ligation into pHSG298. The two gRNAs cloned in pCFD4 vector (gRNA sequences in Table 1) targeted the intronic region between *VGAT* exon 1 and exon 2. The insertion included an ALFA tag as well as a DsRed selection cassette. The whole edit was surrounded by ∼1 kb homology arms.

GluRIIA exon HDR construct in pCR4. The full assembly was inserted by Topo cloning into pCR4. The two gRNAs cloned each in a pCDF3 vector targeted the beginning of exon 12 and the intergenic region between *GluRIIA* and *GluRIIB*. (gRNA sequences in Table 1) The inserted sequence included another version of exon 12 followed by the endogenous 3'UTR and a DsRed selection cassette. The whole edit was surrounded by ∼1 kb homology arms.

7ALFA-Synapsin HDR construct in pVID. The 5' gRNA targeted the intronic region between exon 1 and 2 of *Synapsin*, and the 3' gRNA targeted the intronic region between exon 2 and 3 (gRNA sequences in Table 1). The edit included 7 copies of ALFA tag and an excisable dsRed selection cassette. Upstream of the edit was two gRNAs with their respective scaffold sequence, interspaced by a tRNA sequence (50, 51). The edit was flanked by two 600-700bp Homology arms. The full sequences were synthesized and cloned in a pVID vector by Genscript.

	hTDP43 knock-in		*GluRIIA mutation*		*ALFA-vGat*		*ALFA-Synapsin*	
	5'	3'	5'	3'	5'	3'	5'	3'
gRNA sequence	GGATAATGTAACC GCTGTAC CG**G**	GGCCGGCATCCA CAAGGGTC CG**G**	GCTGAGCGAAAG GGGCGAGC TG**G**	GACTCTAAAGTCT AGACATG TG**G**	GGTGGGTCATTCT CATTA T**G**G	GCCTGCGGAAGA TGGATT T**G**G	GGGCCCCCTTCT GCCAAAGG G**GG**	GGGTATGCTTTGG CA/TTGGC AGG
gRNA Off-target*	0	0	3	0	0	0	1	3
gRNA GC content	45%	70%	70%	40%	44%	56%	70%	55%
gRNA efficiency score**	4.4	5.7	5.6	4.3	5.9	5.8	5.3	4.5
Homology arm size	1082bp	1077bp	1000bp	999bp	897bp	1328bp	604bp	602bp
Cutting distance from edit	1835bp	1895bp	295bp	203bp	18bp	21bp	391bp	5bp
Injection stock	nos-cas9 attp2 BL 78782		nos-cas9 attp2 BL 78782		nos-cas9 attp2 BL 78782		nos-cas9 attp40 BL 78781	
Vector backbone	pCR4		pCR4		pHSG298		pVID (ZsGreen)	
Aberrant vector insertion	18%		91%		67%		50%	
After screening against vector marker							0%	

*http://targetfinder.flycrispr.neuro.brown.edu **https://www.flyrnai.org/evaluateCrispr

gRNA color code : DISTAL PROXIMAL PA**M** **in bold:** nucleotide mutated in donor constrcut

Table 1. Information about the characteristics of the guide RNAs used.

GluRIIC HDR Construct in pVID. The 5' gRNA targeted the middle of exon 10, and the 3' gRNA targeted the intronic region between exon 2 and 3 (gRNA sequences in Table 1). The edit included a Myc tag followed by an excisable dsRed selection cassette. Upstream of the edit was two gRNAs with their respective scaffold sequence, interspaced by a tRNA sequence (50, 51). The edit was flanked by two 600-700bp Homology arms. The full sequences were synthesized and cloned in a pVID vector by Genscript.

pVID HDR vector generation. A *Drosophila* eye negative selection cassette composed of a 3xP3 promoter (48), a ZsGreen fluorescent protein (52, 53) and a SV40 terminator was inserted upstream of a dU6:3 gRNA promotor. Downstream of the promoter a multiple cloning site (MCS) was introduced to allow simple insertion, by restriction digestion and ligation, of the specific gRNAs and the HDR editing template to produce an 'all-in-one' CRISPR genome editing construct. The region including the green negative selection cassette and MCS was synthetized and subcloned into pUC57-mini vector by Genscript. The pVID full sequence is available with Genbank (54) accession: PV446458. Addgene 236635.

Generation of *Drosophila melanogaste*r lines. To generate the transgenic lines, *Drosophila melanogaster* embryos expressing Cas9, under the germ cell specific nos promoter, were injected by either Genetivision (Stafford, TX, USA) or FlyORF (Zürich, Switzerland) with the HDR construct and if required the appropriate gRNA expressing vectors.

Immunoblotting. 20 adult *Drosophila* heads were lysed in RIPA lysis buffer (140 mM NaCl, 1 mM EDTA, 1%Triton X-100, 0.5 mM EGTA, 10 mM Tris-Cl pH 8, 0.1% SDS) by homogenisation with a pestle and sonication for 5 s. The samples were subsequently centrifuged for 1 min at 12 000 x g, the supernatant collected, and the protein concentration measured with a Micro BCA Protein Assay kit (Thermo Scientific). Before loading 30 µg of protein into a 4-12% Tris-Glycine gel (Invitrogen), the samples were mixed with 4x Laemmli buffer and heated for 2 minutes at 85° C. After running the gel, the protein bands were dry transferred to a PVDF membrane (iblot transfer, Invitrogen) which was blocked for 1 h in 5% milk blocking solution. The membrane was then incubated at 4℃ overnight in a mix of blocking solution and primary antibodies. The primary antibodies used were anti-hTDP-43 (H00023435-M01, Novus Biologicals, 1:500), anti-TBPH (custom made, Pierce, 1:500), anti-α-tubulin (T6199, Sigma, 1:10 000). Membranes were incubated with secondary antibodies (goat anti-rabbit, 111-035-003, Jackson ImmunoResearch, 1:10 000) (goat anti-mouse 115-035-003, Jackson ImmunoResearch, 1:10 000) diluted in blocking solution for 1 h. Chemiluminescent substrate (WesternBright Sirius, Advansta) was added to enable imaging. To stain with several consecutive antibodies the membranes were stripped for 10 min at 37℃ (Restore Western Blot Stripping Buffer, Thermo Scientific) and re-blocked.

PCR of region between incorrect insert and endogenous gene. The genomic DNA of all *Drosophila* lines were purified using a commercial kit (DNeasy Blood and Tissue kit, Qiagen). To detect the DNA sequence between the *hTDP-43* gene and the *TBPH* gene in the incorrectly inserted *hTDP-43* knock-in *Drosophila*, we performed a PCR using a proofreading DNA polymerase (Platinum SuperFi II, Invitrogen), a reverse primer in the 5' end of the *hTDP-43* gene (TTCTCATCCTCG-GTCACGCG) and a forward primer in the 3' end of the *TBPH* gene (GGAACACCGGTTGGTCGAAC). The PCR product was run on an agarose gel and the band was purified using a gel extraction kit (GeneJET Gel Extraction Kit, Thermo Scientific) and a reaction clean-up kit (MinElute Reaction cleanup kit, Qiagen). The PCR product was then cloned into a pCR8 vector with pCR™8/GW/TOPO™ TA Cloning Kit (Invitrogen) and sequenced.

Identification of the pCR4 or pHSG298 template vector backbones in the *Drosophila* genome. To identify the presence of the vector backbone in the *Drosophila* genome, we purified the genomic DNA using DNeasy Blood and Tissue kit (Qiagen), and performed a PCR using primers within the lac promoter and pUC ori in the pCR4 or pHSG298 plasmid back- bone (forward: CGCGTAATCTGCTGCTTGCAA, reverse: TGTCGTGCCAGCTGCATTAATG). The resulting band should be 817 bp long in the pCR4 backbone plasmid, and 1001 bp long in the pHSG298 plasmid backbone. For pVID vector backbone, we performed a PCR using primer covering the region between the zsGreen CDS and the dU6:3 promoter (forward: 5'-GACAGACAACTGGGAGCCTAGTTG-3', reverse: 5'-TGAGTAGGTGGCGTTTCATTCTACTC-3'). The resulting band should be 813bp long.

Identifying undesired HDR template vector genomic insertion with pVID. To select only animals where editing has occurred but the template HDR vector has not become incorporated into the genome we selected for DsRed expressing eye fluorescence and subsequently ZsGreen green expressing eye fluorescence. Flies exhibiting ZsGreen eye fluorescence indicted the undesired genomic integration of the pVID vector backbone. Fluorescence screening was performed using a fluorescence stereomicroscope (MC165F Leica) equipped with a GFP and dsRed filter. Images were taken using a fluorescence CCD camera (DFC3000G, Leica).

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-*hTDP-43*	Novus Biologicals	H00023435-M01
Rabbit anti-*TBPH*	This manuscript, Pierce	-
Mouse anti-a-tubulin	Sigma	T6199
Goat anti-rabbit	Jackson ImmunoResearch	111-035-003
Goat anti-mouse	Jackson ImmunoResearch	115-035-003
Chemicals, peptides, and recombinant proteins		
Restore Western Blot Stripping Buffer	Thermo Scientific	21059
WesternBright Sirius	Advansta	K-12043-D20
Critical commercial assays		
DNeasy Blood and Tissue kit	Qiagen	69504
Micro BCA Protein Assay kit	Thermo Scientific	23235
GeneJET Gel Extraction kit	Thermo Scientific	K0691
MinElute Reaction cleanup kit	Qiagen	28204
pCR™8/GW/TOPO™ Cloning Kit	Invitrogen	K250020
Zero Blunt® TOPO® PCR Cloning Kit	Invitrogen	450159
One shot Top10 chemical competent cells	Invitrogen	C404003
Experimental models: Organisms/strains		
Drosophila melanogaster w^{1118}		
Drosophila melanogaster; Incorrectly edited *hTDP-43*;;	This manuscript	-
Drosophila melanogaster; Incorrectly edited mutated *GluRIIA* exon 12;;	This manuscript	-
Drosophila melanogaster; Incorrectly edited ALFA tag-*VGAT*;;	This manuscript	-
Drosophila melanogaster; Incorrectly edited 7ALFA tag-*Synapsin* (pVID);;	This manuscript	-
Drosophila melanogaster; Incorrectly edited GluRIIC-Myc (pVID);;	This manuscript	-
Oligonuclotides		
gRNA 5' *hTDP-43*	Microsynth	GGATAATGTAACCGCTGTAC
gRNA 3' *hTDP-43*	Microsynth	GGCCGGCATCCACAAGGGTC
gRNA 5' *GluRIIA*	Microsynth	CTGAGCGAAAGGGGCGAGC
gRNA 3' *GluRIIA*	Microsynth	GACTCTAAAGTCTAGACATG
gRNA 5' *VGAT*	Microsynth	GGTGGGTCATTCTCATTA
gRNA 3' *VGAT*	Microsynth	GCCTGCGGAAGATGGATT
gRNA 5' *Synapsin*	Genscript	GGGCCCCCTTCTGCCAAAGG
gRNA 3' *Synapsin*	Genscript	GGGTATGCTTTGGCATTGGC
gRNA 5' *GluRIIC*	Genscript	GCACTCCAATCTCCAGTGTC
gRNA 3' *GluRIIC*	Genscript	GGCACCCTTCCCCCCGCTGG
Primer *TBPH* 3' forward	Microsynth	GGAACACCGGTTGGTCGAAC
Primer *hTDP-43* 5' reverse	Microsynth	TTCTCATCCTCGGTCACGCG
Primer pUC forward	Microsynth	CGCGTAATCTGCTGCTTGCAA
Primer ZsGreen forward	Microsynth	GACAGACAACTGGGAGCCTAGTTG
Primer dU6:3 Re	Microsynth	TGAGTAGGTGGCGTTTCATTCTACTC
Recombinant DNA		
pVID	Addgene	236635
pScarlessHD-DsRed	Addgene	64703