JCB

CRISPR-Cas12a-assisted PCR tagging of mammalian genes

Julia Fueller, Konrad Herbst, Matthias Meurer, Krisztina Gubicza, Bahtiyar Kurtulmus, Julia Knopf, Daniel Kirrmaier, Benjamin Buchmuller, Gislene Pereira, Maurius Lemberg, and Michael Knop

Corresponding Author(s): Michael Knop, University of Heidelberg

Review Timeline:

Submission Date:	2019-10-30
Editorial Decision:	2019-12-09
Revision Received:	2020-02-10
Editorial Decision:	2020-02-20
Revision Received:	2020-03-10

Monitoring Editor: Bas van Steensel

Scientific Editor: Marie Anne O'Donnell

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

DOI: https://doi.org/10.1083/jcb.201910210

December 9, 2019

Re: JCB manuscript #201910210

Prof. Michael Knop
University of Heidelberg
ZMBH
Im Neuenheimer Feld 282
Heidelberg 69120
Germany

Dear Prof. Knop,

Thank you for submitting your manuscript entitled "CRISPR-Cas12a-assisted PCR tagging of mammalian genes". The manuscript was assessed by expert reviewers, whose comments are appended to this letter. We invite you to submit a revision if you can address the reviewers' key concerns, as outlined here.

You will see that Reviewer #1 is very positive and the comments of Reviewer #2 are constructive and focused. Addressing these points is necessary to bolster the study's main claims and they appear mostly addressable by providing some additional experimental data or by edits of the manuscript. In addition, it was recently reported that homology-based editing can be boosted substantially by treatment of cells with the NHEJ inhibitor M3814 (https://academic.oup.com/nar/article/47/19/e116/5545003). This compound would also suppress NHEJ-mediated concatemerization of the template as described in the manuscript. In the Discussion, testing NHEJ-inhibitors is mentioned as "data not shown" - but was M3814 tested? This compound caused a lot of excitement at a recent CRISPR meeting at Cold Spring Harbor. If the tagging efficiency is enhanced by M3814, then this would strengthen the manuscript further. I recommend to test and report this, perhaps together with any other NHEJ inhibitors that were tested. This would then also address the issue that JCB discourages "data not shown" statements.

While you are revising your manuscript, please also attend to the following editorial points to help expedite the publication of your manuscript. Please direct any editorial questions to the journal office.

GENERAL GUIDELINES:

Text limits: Character count for an Tools is < 40,000, not including spaces. Count includes title page, abstract, introduction, results, discussion, acknowledgments, and figure legends. Count does not include materials and methods, references, tables, or supplemental legends.

Figures: Toolss may have up to 10 main text figures. Figures must be prepared according to the policies outlined in our Instructions to Authors, under Data Presentation, http://jcb.rupress.org/site/misc/ifora.xhtml. All figures in accepted manuscripts will be screened prior to publication.

***IMPORTANT: It is JCB policy that if requested, original data images must be made available.

Failure to provide original images upon request will result in unavoidable delays in publication. Please ensure that you have access to all original microscopy and blot data images before submitting your revision.***

Supplemental information: There are strict limits on the allowable amount of supplemental data. Toolss may have up to 5 supplemental figures. Up to 10 supplemental videos or flash animations are allowed. A summary of all supplemental material should appear at the end of the Materials and methods section.

The typical timeframe for revisions is three months; if submitted within this timeframe, novelty will not be reassessed at the final decision. Please note that papers are generally considered through only one revision cycle, so any revised manuscript will likely be either accepted or rejected.

When submitting the revision, please include a cover letter addressing the reviewers' comments point by point. Please also highlight all changes in the text of the manuscript.

We hope that the comments below will prove constructive as your work progresses. We would be happy to discuss them further once you've had a chance to consider the points raised in this letter.

Thank you for this interesting contribution to Journal of Cell Biology. You can contact us at the journal office with any questions, cellbio@rockefeller.edu or call (212) 327-8588.

Sincerely,

Bas van Steensel, Ph.D.
Monitoring Editor

Marie Anne O'Donnell, Ph.D.
Scientific Editor

Journal of Cell Biology

Reviewer #1 (Comments to the Authors (Required)):

Fueller et al. describes a gene tagging method in mammalian cells that is based on the Cas12a coupled with a PCR cassette that serve both as the DNA repair donor and the template for expressing the guide RNA. While many methods exist for endogenous gene tagging, this method is unique as it simplifies the components required for tagging and as such can reduce the cost for the geneartion of libraries of tagged cell lines. In addition, it is apprent that the authors have made a significant effort to make their method very accesible to the scientific community. The data is solid and well described and there is an investigation into the underlying causes of outlier effects (e.g. the diffused cytoplasmic fluorescence). In general I would recommend publication with minor comments below:

(1) It is hard to understand the method from figure 1, for example when you get to 1c it is hard to remember what each color stands for, I would try to re design figures 1a-c in order to make the method more intuitive.
(2) How likely is it for a tagged cell line of some gene X to have a KO in the non tagged allele?

Reviewer #2 (Comments to the Authors (Required)):

Fueller et al. describe a method for tagging mammalian genes using PCR-generated linear dsDNA donors. Their work is unique with regards to the linking in cis of the crRNA expression cassette for Cas12a. The concept is an extension of their work on CRISPR-Cas12a-assisted gene tagging in yeast (Buchmuller, Nat. comm. 2019). The authors created a series of templates for C-terminal PCR tagging and a bio-informatic pipeline to facilitate the adoption of their method. While this approach is conceptually similar to several published studies, it distinguishes itself by clearly exposing its limitations. It appears that the most useful aspect of the method lies in the analysis of enriched populations of targeted cells, which will save time if researchers are willing to compromise and work without "perfectly characterized cell lines with exactly the intended genomic modification". With that in mind, it is worth considering if targeted integration of a tagged cDNA at a genomic safe harbor such as AAVS1 would not be easier, faster, more efficient, and as informative (please see the last paragraph of this review).

Specific points:

1. Abstract: should state that the method is; (i) designed for c-term tagging, (ii) leaves ectopic DNA elements (e.g. polyA, crRNA expression cassette, selection markers) at the target loci [or mention that it is not "seamless" integration of the tag], (iii) uses linear dsDNA donors, (iv) a web-based interface has been created to facilitate design.

2. Intro: The authors state that "The protospacer-adjacent motif (PAM) of Cas12a is T-rich in comparison to the G-rich PAM of Cas9 which identifies more possible target sequences near STOP codons of mammalian cells (Buchmuller et al., 2019)." This analysis was performed using a TYN PAM which is not representative of the current experimental setup. This sentence should be adapted accordingly.

3. Intro: The authors state that "This makes chromosomal tagging rather cumbersome, time consuming and costly." Considering the cost of the long ~110-120bp M1/M2 oligos (especially if they have to be modified) this sentence should be adapted. Specifically, can the authors compare the cost of synthesis for their oligos to ordering of a gBlock (linear dsDNA fragment with short homology arms and the tag) and cloning oligos into a single expression vector for Cas12a and its crRNA cassette? In other words, it is unclear to the reviewer that the proposed experimental design is less "cumbersome, time consuming and costly" than more 'traditional' alternatives. It also does not circumvent the need to test and identify active Cas12a crRNA. Thus, several donors (M2 oligos) may have to be constructed before finding an appropriate configuration.

4. Results: The authors focus on highly expressed genes that may be easier to target due to increased DNA accessibility. This possibility should be discussed as several evidence indicate that DNA repair can be influenced by chromatin "status".

5. Figure 2 describes an important result and exposes a limitation of the method e.g. rearrangement of the linear dsDNA donors in cells. However, there is extensive literature describing the behavior of linear DNA in cells and it should have been expected. While the authors should state the main conclusion from this thorough analysis, it may be better to present these data in a supplemental figure (see next comment).

6. Figure 3 describes parameters influencing tagging efficiency, but this analysis is limited to HEK293 and few target genes. While they align with previously published work e.g. Orlando et al. PMID: 20530528, these data are not "enabling" and would also be more appropriate for the supplemental material. It should also be stated that "donor homology length is relatively unimportant when HDR copies a small (17 bp) insert into the chromosome yet becomes more important when transgene-size (~1.5 kb) segments are copied (Orlando et al)" since this is directly related to the current work.

*Regarding points 5 and 6. Of course, the reviewer does not impose these changes to the manuscript to the authors. However, the reviewer feels that these sections distract the reader from the main points of the manuscript. Collectively, the data presented in F2 and F3 could be summarized in 1-2 sentences at the end of the description of F1. Importantly, please clearly state if protected primers (e.g. 5S biotin used in F4) is the preferred choice for M1 and M2.

7. Figure 4. The authors write: "the smaller and more abundant bands corresponding to the size of the junctions formed by HR tag insertion, and larger and less abundant bands corresponding to the size expected from fragment insertions by c-NHEJ". The abundance of both bands is influenced by PCR "efficiency" and should not be used to "qualify/imply" the ratio of HDR/NHEJ-mediated integration.

8. Figure 4D. The authors should also quantify the ratio of "perfect" vs "erroneous" integrations at the other end of the break (between the tag and the 3'UTR). It is very likely that a fraction of these repair events will also be "erroneous" in the population. This analysis will prevent any "inflation" of the % of "perfect" integrations. At the minimum, it should be stated that integration of the tag could occur on one side of the break via HDR and on the other side via NHEJ (e.g. https://www.biorxiv.org/content/10.1101/841098v1).

9. Figure 4F. These data indicate that the activity of Cas12a is quite limited (low % indels) and suggests that delivery of the crRNA expression cassette as a PCR product limits Cas12a efficiency in this system. It would be useful to compare the % indels generated when crRNAs are provided as a PCR product versus a plasmid vector (same molar amounts). If it is the case, then a more "traditional" way of providing the crRNA cassette should be considered/mentioned by the authors.

10. Figure 5. The authors observe that "correctly tagged clones contain frequently integrated concatemers" at the other alleles. This is a major limitation of the method and should be discussed since it leads to high levels of "on-target" heterogeneity. What is the % indels at the non-targeted alleles (no-integration of the donor) in the selected populations?

11. Figure 6. Tagging levels appear to be marginal for most targets in unselected cell lines.

12. Figure 7. The authors should clearly state which PAMs have been used to make part B. For example, have they included Tier 1-2-3 PAMs for enAsCas12a? The statement "we calculated that close to 100% of all human ORFs (Fig. 7b) are amenable for C-terminal PCR tagging." should be put into perspective.

13. Table 1. Please include a column with Addgene plasmid #.

14. Discussion. In order to minimize the impact of using an heterologous polyA, and the concomitant insertion of a complete crRNA expression cassette at the 3'UTR of tagged genes, the authors state "While for mammalian cells no global data set about the regulatory impact of the 3'-UTR on

gene expression is available, data from yeast, where seamless tagging was compared with tagging using a generic 3'-UTR, demonstrated that only about 11% of the genes were impacted in their expression more than 2-fold." The authors should not fail to mention the existence and the role of miRNAs in mammalian systems.

For the reasons mentioned above, the reviewer believes that the authors failed to provide sufficient evidence to support the following sentences: "In conclusion, PCR-mediated gene tagging has the potential to impact how research is done in an entire field, not only because of the simplicity of the method, but also because the required reagents are easy to handle, cost effective, and freely exchangeable. Moreover, PCR tagging is commonly quicker than the construction of a plasmid for transient transfection, while simultaneously alleviating the danger of studying overexpression artifacts." In that regard, this main conclusion is an overstatement.

The reviewer recognizes the value of this work. However, there is a disconnect between what is claimed and the actual data. In absence of direct comparison with previously developed methods to achieve the same outcome it is difficult to conclude that the proposed approach is more accurate, simpler, more cost effective, and faster. It may well be that focusing on the description of the method and of its limitations while avoiding claims regarding the superiority of the approach would better serve this manuscript. There was a lot of work put into this manuscript and it should not remain unpublished.

Line numbers are provided that refer to changes in the manuscript, highlighted in the attached pdf file: MammaTag_v011_Rev_v06_incl TrackChanges)

Comments from Editor:
You will see that Reviewer #1 is very positive and the comments of Reviewer #2 are constructive and focused. Addressing these points is necessary to bolster the study's main claims and they appear mostly addressable by providing some additional experimental data or by edits of the manuscript. In addition, it was recently reported that homology-based editing can be boosted substantially by treatment of cells with the NHEJ inhibitor M3814 (https://academic.oup.com/nar/article/47/19/e116/5545003). This compound would also suppress NHEJ-mediated concatemerization of the template as described in the manuscript. In the Discussion, testing NHEJ-inhibitors is mentioned as "data not shown" - but was M3814 tested? This compound caused a lot of excitement at a recent CRISPR meeting at Cold Spring Harbor. If the tagging efficiency is enhanced by M3814, then this would strengthen the manuscript further. I recommend to test and report this, perhaps together with any other NHEJ inhibitors that were tested. This would then also address the issue that JCB discourages "data not shown" statements.

Answer: Thank you for raising this point for discussion and further clarification. We tested Scr7 und 3-Aminobenzamid. Both did not show any effect in terms of increased tagging efficiency. Based on your suggestion, we now also tested M3814, and found that it did not enhance tagging efficiency either. For the purpose of this point-by-point response, we have included some of the data we produced during assessment of the influence of M3814 (see Figure below), but because of various reasons explained below, we think that we should not show the data in the paper, and we will therefore also remove the statement “(data not shown)” from the manuscript.

In the report from the Pääbo lab (PMID: 31392986), M3814 was effective to enhance the repair of CRISPR-induced DSBs and DNA nicks. Notably, this study used short ssDNA oligonucleotides as repair templates in order to induce single/few nucleotide substitutions. The diverse nature of the donor templates of their and our study (short ssDNA versus long dsDNA) potentially leads to differential utilization of homologous repair pathways. This issue, if true, makes the findings of their and our study incomparable. Although the identification of the underlying pathways would undoubtedly be interesting, further investigations will be needed. In addition, we observed that M3814 is very toxic to HEK293T cells, which required us to use much lower concentrations. We contacted an author from the Pääbo study and he confirmed toxicity of the inhibitor for the related cell line HEK293. This toxicity was not seen as prominent in the specific cell lines they used in their published study.

Furthermore, an additional survey of the Scr7 literature identified a paper that challenged the specificity of Scr7 for DNA ligase IV, suggesting an inhibitory mechanism different from c-NHEJ ablation (PMID: 27235626) and in extension questions the value of this experiment.

From these assessments we reasoned that, in order to obtain quantitative and useful information about the impact of different ways to inhibit different non-HR pathways, a much more thorough study using our tagging strategy and other genome editing methods in comparison would be needed, in order to provide new and helpful insights. We would like to argue that this is out of the scope of this study.

Based on these reasons we have removed the statement about NHEJ along with “(data not shown)”.

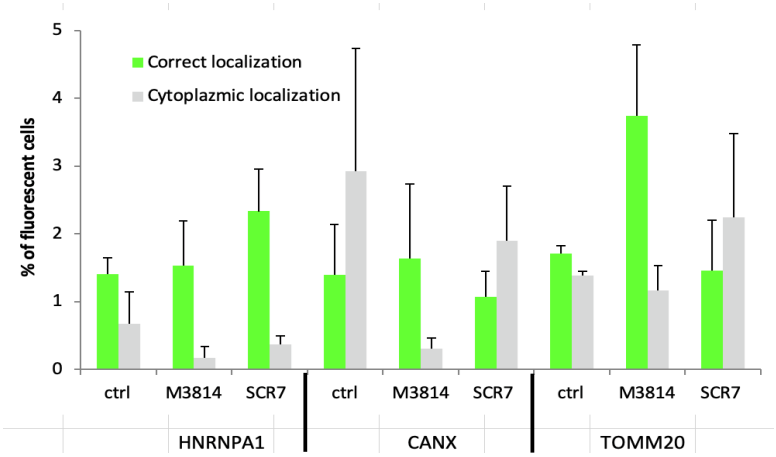

Figure: GFP-Tagging of three genes as a function of treatment of the cells with 1 µM Scr7 or 0.2 µM M3814. To account for strong toxic effects of M3814 the transfection was conducted for 4 hours only (instead of 24 hours) thus leading to lower transfection efficiency. Data from one experiment (out of 3 experiments with different transfection times) is shown. Error bars: SD of 3 technical replicas.

Modification in the manuscript:
Line 617-621: Further improvements of the tagging efficiency might be possible, ~~While we found that inhibitors of c-NHEJ did not exhibit a positive impact on the integration efficiency (data not shown), it might be,~~ i.e. by targeting the repair template to the CRISPR endonuclease cut site (Roy et al., 2018), or by using Cas12a variants that are only active in S/G2 phase of the cell cycle ~~in order to improve tagging efficiency and fidelity further~~ (Smirnikhina et al., 2018).

Reviewer #1 (Comments to the Authors (Required)):

Fueller et al. describes a gene tagging method in mammalian cells that is based on the Cas12a coupled with a PCR cassette that serve both as the DNA repair donor and the template for expressing the guide RNA. While many methods exist for endogenous gene tagging, this method is unique as it simplifies the components required for tagging and as such can reduce the cost for the geneartion of libraries of tagged cell lines. In addition, it is apprent that the authors have made a significant effort to make their method very accesible to the scientific community. The data is solid and well described and there is an investigation into the underlying causes of outlier effects (e.g. the diffused cytoplasmic fluorescence). In general I would recommend publication with minor comments below:

(1) It is hard to understand the method from figure 1, for example when you get to 1c it is hard to remember what each color stands for, I would try to re design figures 1a-c in order to make the method more intuitive.

Answer:
We thank the reviewer for pointing this out. We have redesigned figure panels 1a-c and condensed the information into two panels (new Fig. 1a and b). We believe that these panels are better suited to allow that readers can gain intuitively access to the main elements of the method. We merged the information of the previous panels a and b into one new panel (new Fig. 1a), by using a different visual strategy, that we believe is more intuitive and easier to grasp by the readers. Legends and text were updated accordingly.

Modification in the manuscript:
Modified Figure panels 1a-b and updated corresponding figure legends.

(2) How likely is it for a tagged cell line of some gene X to have a KO in the non tagged allele?

Answer: We thank the reviewer to pointing this out. In all genome editing methods the possibility exists that homologous chromosomes are repaired using different/alternative pathways, and only one incorporates the desired change. It is therefore possible that a cell has received a correct insertion on one chromosome and has repaired the second DSB on the homologous chromosome using for example c-NHEJ. In case this repair introduces a DNA sequence alteration (indel) it would affect the ORF near its 3'end, where the tag should be inserted, either before or after the stop codon. We are aware of one report from one of the users of our method who observed the untagged protein was no longer observed, which could be explained by NMD of the damaged mRNA. However, since we do not have extensive data about many such cases we cannot make a general statement in order to answer this question. While we already show in In Fig 5b that in about 1/3 of the cases also the homologous chromosome receives a functional tag insertion, we assume that the efficiency/frequency of the other allele being modified in a way leading to the abrogation of protein expression depends on various parameters. We think that making a general quantitative statement should therefore be avoided. Furthermore, HEK293T cells are also polyploid for many chromosomes, which complicates the discussion even further. Given that we also discuss the consequences of DSB repair by NHEJ (Fig 4f, g) we do not think that this point needs more discussions.

Modification in the manuscript:
none

Reviewer #2 (Comments to the Authors (Required)):

Fueller et al. describe a method for tagging mammalian genes using PCR-generated linear dsDNA donors. Their work is unique with regards to the linking in cis of the crRNA expression cassette for Cas12a. The concept is an extension of their work on CRISPR-Cas12a-assisted gene tagging in yeast (Buchmuller, Nat. comm. 2019). The authors created a series of templates for C-terminal PCR tagging and a bio-informatic pipeline to facilitate the adoption of their method. While this approach is conceptually similar to several published studies, it distinguishes itself by clearly exposing its limitations. It appears that the most useful aspect of the method lies in the analysis of enriched populations of targeted cells, which will save time if researchers are willing to compromise and work without "perfectly characterized cell lines with exactly the intended genomic modification". With that in mind, it is worth considering if targeted integration of a tagged cDNA at a genomic safe harbor such as AAVS1 would not be easier, faster, more efficient, and as informative (please see the last paragraph of this review).

Indeed, safe harbor integration of tagged genes is a valid and frequently used method. Still, for safe harbor integration a cloning step is needed, which is more time consuming, and which makes multiplexing more difficult. And, PCR tagging relies on endogenous tagging, which reduces the risk of artefacts associated with ectopic expression using heterologous promoters, as used in safe harbor expression.

Specific points:

1. Abstract: should state that the method is; (i) designed for c-term tagging, (ii) leaves ectopic DNA elements (e.g. polyA, crRNA expression cassette, selection markers) at the target loci [or mention that it is not "seamless" integration of the tag], (iii) uses linear dsDNA donors, (iv) a web-based interface has been created to facilitate design.

Answer: We have modified the abstract to contain this information.

Modification in the manuscript: Abstract changed accordingly

2. Intro: The authors state that "The protospacer-adjacent motif (PAM) of Cas12a is T-rich in comparison to the G-rich PAM of Cas9 which identifies more possible target sequences near STOP codons of mammalian cells (Buchmuller et al., 2019)." This analysis was performed using a TYN PAM which is not representative of the current experimental setup. This sentence should be adapted accordingly.

Answer: We apologize for the confusion. We now removed this sentence (underlined above), since this information is of no particular value for the reader.

Modification in the manuscript:
Lines 50-52: Removed respective sentence from introduction.

3. Intro: The authors state that "This makes chromosomal tagging rather cumbersome, time consuming and costly." Considering the cost of the long ~110-120bp M1/M2 oligos (especially if they have to be modified) this sentence should be adapted. Specifically, can the authors compare the cost of synthesis for their oligos to ordering of a gBlock (linear dsDNA fragment with short homology arms and the tag) and cloning oligos into a single expression vector for Cas12a and its crRNA cassette? In other words, it is unclear to the reviewer that the proposed experimental design is less "cumbersome, time consuming and costly" than more 'traditional' alternatives. It also does not circumvent the need to test and identify active Cas12a crRNA. Thus, several donors (M2 oligos) may have to be constructed before finding an appropriate configuration.

Answer: We thank the reviewer to point this out. Indeed, we probably could have compared the costs for the DNAs for transfection using different methods. However, these are subject to changes, and the material costs do not reflect the true costs, which would require to also account for the time used to produce PCR products versus cloned fragments. Therefore, we propose an alternative solution to this discussion by simply removing this sentence and to leave it to the reader to judge these aspects from their own point of view.

Modification in the manuscript
Removed respective sentence from introduction.
Lines 56-66: Removed respective sentence from introduction.

4. Results: The authors focus on highly expressed genes that may be easier to target due to increased DNA accessibility. This possibility should be discussed as several evidence indicate that DNA repair can be influenced by chromatin "status".

Indeed, the reviewer raises an important point. By using a visual output to estimate the transfection success we obviously focused on genes with a ‘decent’ expression level. Since in yeast we did not find a correlation between Cas12a mediated tagging efficiency and expression levels (Buchmuller et al., 2019), we somehow haven’t really considered this point. But obviously the situation in mammalian cells might be different. To account for this, we extended the discussion.

Modification in the manuscript:
We introduced the following sentences:
Lines 556-561: Tagging efficiency might be influenced by various factors including chromatin structure and expression levels. Our choice of relatively highly expressed genes as convenient reporters to validate and investigate the method might bias the tagging efficiency. While genome wide analysis of tagging efficiency for Cas12a tagging using yeast did not reveal a correlation of expression levels and tagging efficiency (Buchmuller et al., 2019), further experiments will be needed to validate whether this is also the case in mammalian cells.

5. Figure 2 describes an important result and exposes a limitation of the method e.g. rearrangement of the linear dsDNA donors in cells. However, there is extensive literature describing the behavior of linear DNA in cells and it should have been expected. While the authors should state the main conclusion from this thorough analysis, it may be better to present these data in a supplemental figure (see next comment).

Answer Points 5. + 6.: We agree with the reviewer that our study in the end only confirms previous work, i.e. by the Capecchi lab (Refs Folger et al., 1982,1985) where the behavior of linear DNA injected into mammalian cells has been thoroughly investigated. However, these previous studies are not necessarily known to the current generation of researchers working with mammalian cells and we would therefore prefer to showcase this aspect in the context of our

tagging strategy. In addition, we consider this aspect to be important for the understanding of our workflow. Hence, we prefer to show the results in the main part of our paper.

6. Figure 3 describes parameters influencing tagging efficiency, but this analysis is limited to HEK293 and few target genes. While they align with previously published work e.g. Orlando et al. PMID: 20530528, these data are not "enabling" and would also be more appropriate for the supplemental material. It should also be stated that "donor homology length is relatively unimportant when HDR copies a small (17 bp) insert into the chromosome yet becomes more important when transgene-size (~1.5 kb) segments are copied (Orlando et al)" since this is directly related to the current work.

*Regarding points 5 and 6. Of course, the reviewer does not impose these changes to the manuscript to the authors. However, the reviewer feels that these sections distract the reader from the main points of the manuscript. Collectively, the data presented in F2 and F3 could be summarized in 1-2 sentences at the end of the description of F1. Importantly, please clearly state if protected primers (e.g. 5S biotin used in F4) is the preferred choice for M1 and M2.

We thank the reviewer for noting that the importance of modified primers is not yet apparent. We modified the discussion accordingly to make the suggestion clearer.

Modification in the manuscript:
We introduced the following sentence to the discussion:
Lines 615-616: We recommend the use of chemically modified M1 and M2 primers (e.g. with 5S and biotin) as we noticed considerable enhancement in tagging fidelity.

7. Figure 4. The authors write: "the smaller and more abundant bands corresponding to the size of the junctions formed by HR tag insertion, and larger and less abundant bands corresponding to the size expected from fragment insertions by c-NHEJ". The abundance of both bands is influenced by PCR "efficiency" and should not be used to "qualify/imply" the ratio of HDR/NHEJ-mediated integration.

Answer: We agree that the PCR efficiency can differ between fragments. However, in our case the fragments are very similar, since they possess the same primer binding sites and in large parts the same sequence. In our experience, when using processive polymerases and carefully optimized PCR conditions (which we used), PCR can be used to approximate the relative abundance of the different templates and make at least a qualitative statement. Moreover, since we do not correct the signal intensity for fragment length, we overestimate the abundance of the longer fragment, which works in our favor. However, to make the reader more aware about these considerations, we modified this paragraph:

Modification in the manuscript:
Lines 317-323: We used PCR to amplify the insertion junction between the 3' of the ORF and the inserted tag. This yielded two distinct amplicon population. The shorter bands correspond in their size to the junctions formed by HR tag, and the longer to the size expected from fragment insertions by c-NHEJ (Fig. 4a-c). Despite that PCR of not fully identical fragments can differ in efficiency, the results suggest the more insertion junction formed by HR are present in the population. Illumina dye sequencing of the shorter bands […].

8. Figure 4D. The authors should also quantify the ratio of "perfect" vs "erroneous" integrations at the other end of the break (between the tag and the 3'UTR). It is very likely that a fraction of these repair events will also be "erroneous" in the population. This analysis will prevent any "inflation" of the % of "perfect" integrations. At the minimum, it should be stated that integration of the tag could occur on one side of the break via HDR and on the other side via NHEJ (e.g. https://www.biorxiv.org/content/10.1101/841098v1).

Answer: We agree with the reviewer. In fact, we mention this point in the Discussion, where we wrote:

PCR tagging does not generate seamlessly integrated tag, since it is accompanied by a generic transcription termination site that replaces the endogenous 3'-UTR. This actually bears the advantage that it reduces the errors associated with tag insertion, since an erroneous insertion downstream of the PCR cassette, i.e. caused by c-NHEJ instead of homologous recombination will only affect the 3'-UTR of the gene, which is not used for the tagged allele.

Modification in the manuscript:
none

9. Figure 4F. These data indicate that the activity of Cas12a is quite limited (low % indels) and suggests that delivery of the crRNA expression cassette as a PCR product limits Cas12a efficiency in this system. It would be useful to compare the % indels generated when crRNAs are provided as a PCR product versus a plasmid vector (same molar amounts). If it is the case, then a more "traditional" way of providing the crRNA cassette should be considered/mentioned by the authors.

Answer: We did do intend to make the claim that our method is more efficient than other gene tagging approaches using e.g. plasmid vectors for crRNA delivery. Our point is that there is no other method (including transient transfections) that is as simple as ours to quickly generate cells that endogenously express a tagged gene.

Modification in the manuscript:
none

10. Figure 5. The authors observe that "correctly tagged clones contain frequently integrated concatemers" at the other alleles. This is a major limitation of the method and should be discussed since it leads to high levels of "on-target" heterogeneity.

Answer: We apologize that our phrasing here was not optimal and must have led to a misunderstanding. We in fact have no evidence that "the other alleles" would contain frequently integrated concatemers. Instead, our data suggest that a correctly tagged locus frequently contains integrated concatemers (instead of only one copy). We then discuss that "additional copies are unlikely to interfere with the tagged gene since they are insulated from the inserted tag by a proper transcription terminator." Therefore, this is clearly not a major limitation of the method. To make this clearer, we modified the text slightly.

Modification in the manuscript:
Line 394-395: Therefore, it appears that correctly tagged clones contain frequently integrated concatemers at the tagged locus, as also predicted from previous work (Folger et al., 1985; 1982).

What is the % indels at the non-targeted alleles (no-integration of the donor) in the selected populations?

Answer: See Answer to Point (2) of Reviewer #1.

11. Figure 6. Tagging levels appear to be marginal for most targets in unselected cell lines.

Answer: Many cell lines are notorious for low transfection efficiencies. However, we did not normalize to the transfection efficiency.

Modification in the manuscript: To account for this, we modified the following paragraph:

Line 431-442: So far, we have described and characterized mammalian PCR tagging as a robust workflow for chromosomal tagging in HEK293T and HEK293 cells. To challenge the general applicability of PCR tagging, we tested additional human but also murine cell lines to target genes already tagged successfully in our initial experiments. In each cell line we identified for most

genes cells that showed correctly localized green fluorescence. However, we note that for some of these cell lines transfection efficiency was in the lower range so that we observed a tagging frequency of 0.2 to 5% (Fig. 6a-d). Examples of tagged murine myoblast (C2C12) cells are shown in Fig. S6a. For HeLa cells that also do provide only moderate transfection levels, we additionally subjected the cells to selection, and found up to 40% of cells exhibiting the correct localization (Fig. S6b). In conclusion, these results demonstrate that PCR tagging works for different mammalian cell lines and species, including differentiated and stem cells whereby combining transfection with selection vastly increases tagging efficiency.

12. Figure 7. The authors should clearly state which PAMs have been used to make part B. For example, have they included Tier 1-2-3 PAMs for enAsCas12a? The statement "we calculated that close to 100% of all human ORFs (Fig. 7b) are amenable for C-terminal PCR tagging." should be put into perspective.

Modification in the manuscript:
We provide a new plot and we list the PAM sites used for the calculation in the legend. For enAsCas12a we used only Tier 1 and 2 PAM sites.

13. Table 1. Please include a column with Addgene plasmid #.

Answer: We are happy to follow the reviewer's suggestion and include this information but decided to place it in the Supplement to not overload Table 1.

Modification in the manuscript:
Information on Addgene plasmid identifier is now included in Supplementary Table 2 which is now also referred to in the description of Table 1.

14. Discussion. In order to minimize the impact of using an heterologous polyA, and the concomitant insertion of a complete crRNA expression cassette at the 3'UTR of tagged genes, the authors state "While for mammalian cells no global data set about the regulatory impact of the 3'-UTR on gene expression is available, data from yeast, where seamless tagging was compared with tagging using a generic 3'-UTR, demonstrated that only about 11% of the genes were impacted in their expression more than 2-fold." The authors should not fail to mention the existence and the role of miRNAs in mammalian systems.

Answer: We are happy to follow the reviewer's suggestion.

Modification in the manuscript:
Lines 590-591: Obviously, this constitutes a compromise, and includes the possibility that important gene regulatory sequences are omitted from the tagged gene, e.g. miRNA binding sites, RNA targeting motifs or sequences regulating mRNA stability. While for mammalian cells no global data set about the regulatory impact of the 3'-UTR on gene expression is available, data from yeast, where seamless tagging was compared with tagging using a generic 3'-UTR, demonstrated that only about 11% of the genes were impacted in their expression more than 2-fold (Meurer et al., 2018).

For the reasons mentioned above, the reviewer believes that the authors failed to provide sufficient evidence to support the following sentences: "In conclusion, PCR-mediated gene tagging has the potential to impact how research is done in an entire field, not only because of the simplicity of the method, but also because the required reagents are easy to handle, cost effective, and freely exchangeable. Moreover, PCR tagging is commonly quicker than the construction of a plasmid for transient transfection, while simultaneously alleviating the danger of studying overexpression artifacts." In that regard, this main conclusion is an overstatement.

Answer: We agree with the reviewer. We can only extrapolate from yeast, where PCR tagging indeed changed the entire field, but the mammalian field is different and the situation much more complex. We therefore rephrased this paragraph:

Line 625-630: In conclusion, PCR-mediated C-terminal gene tagging is a simple, non-commercial, easily adoptable method to exploratively study protein localization or to explore other functional aspects using endogenous-level expression. It is simple to design the oligos (www.pcr-tagging.com) and open access to all other resources is granted (via Addgene or colleagues), and reagents can be freely exchanged. We believe that for many applications PCR tagging is quicker than the construction of a plasmid for transient transfection or exogenous chromosomal integration.
[…]

February 20, 2020

RE: JCB Manuscript #201910210R

Prof. Michael Knop
University of Heidelberg
ZMBH
Im Neuenheimer Feld 282
Heidelberg 69120
Germany

Dear Prof. Knop:

Thank you for submitting your revised manuscript entitled "CRISPR-Cas12a-assisted PCR tagging of mammalian genes". We would be happy to publish your paper in JCB pending final revisions necessary to meet our formatting guidelines (see details below).

- Figure 1c, e: please indicate in the figure legend the number of independent experiments that the data are based on.

- Figure 5a,b; 6a,b; 7c: data are from one representative experiment. Please include the corresponding other experiments that were not shown (or as mean +/- SD across all experiments) as supplementary figures.

- Provide main and supplementary text as separate, editable .doc or .docx files
- Provide figures as separate, editable files according to the instructions for authors on JCB's website, paying particular attention to the guidelines for preparing images at sufficient resolution for screening and production
- Supplementary text is for Supplementary materials legends only, please move other information to Materials and Methods section of main text
- There is a limit of five supplementary figures there are currently six, please combine two where most appropriate
- Add scale bars to insets Fig 1D, 5C,
- Add MW markers to blots in Fig 8B,
- Provide tables as excel files
- Add a paragraph after the Materials and Methods section briefly summarizing the online supplementary materials

***Thank you for your attention to these final processing requirements. Please revise and format the manuscript and upload materials within 7 days or let us know if you need more time. ***

To avoid unnecessary delays in the acceptance and publication of your paper, please read the following information carefully.

A. MANUSCRIPT ORGANIZATION AND FORMATTING:

Full guidelines are available on our Instructions for Authors page, http://jcb.rupress.org/submission-guidelines#revised. **Submission of a paper that does not conform to JCB guidelines will delay the acceptance of your manuscript.**

B. FINAL FILES:

Please upload the following materials to our online submission system. These items are required prior to acceptance. If you have any questions, contact JCB's Managing Editor, Lindsey Hollander (lhollander@rockefeller.edu).

-- An editable version of the final text (.DOC or .DOCX) is needed for copyediting (no PDFs).

-- High-resolution figure and video files: See our detailed guidelines for preparing your production-ready images, http://jcb.rupress.org/fig-vid-guidelines.

-- Cover images: If you have any striking images related to this story, we would be happy to consider them for inclusion on the journal cover. Submitted images may also be chosen for highlighting on the journal table of contents or JCB homepage carousel. Images should be uploaded as TIFF or EPS files and must be at least 300 dpi resolution.

It is JCB policy that if requested, original data images must be made available to the editors. Failure to provide original images upon request will result in unavoidable delays in publication. Please ensure that you have access to all original data images prior to final submission.

The license to publish form must be signed before your manuscript can be sent to production. A link to the electronic license to publish form will be sent to the corresponding author only. Please take a moment to check your funder requirements before choosing the appropriate license.

Please contact the journal office with any questions, cellbio@rockefeller.edu or call (212) 327-8588.

Thank you for this interesting contribution, we look forward to publishing your paper in Journal of Cell Biology.

Sincerely,

Bas van Steensel, Ph.D.
Monitoring Editor

Marie Anne O'Donnell, Ph.D.
Scientific Editor

Journal of Cell Biology
