EMBO reports

Unexpected gene activation following CRISPR-Cas9-mediated genome editing

Anna G Manjón, Simon Linder, Hans Teunissen, Anoek Friskes, Wilbert Zwart, Elzo de Wit, and René H Medema

Corresponding author(s): Rene Medima, Oncode Institute, the Netherlands Cancer Institute (r.medema@nki.nl)

Review Timeline:

Transfer from Review Commons:	28th Aug 21
Editorial Decision:	3rd Sep 21
Revision Received:	27th Oct 21
Editorial Decision:	26th Nov 21
Revision Received:	29th Nov 21
Accepted:	30th Nov 21

Review COMMONS

Editor: Esther Schnapp

Transaction Report: This manuscript was transferred to EMBO Reports following peer review at Review Commons.

(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. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

Dear Rene,

Thank you for the submission of your manuscript with referee reports from Review Commons. I think your study would be a nice contribution to our research section, and your proposed revision plan sounds good.

I would thus like to invite you to revise your manuscript with the understanding that the referee concerns must be fully addressed and their suggestions taken on board. Please address all referee concerns in a complete point-by-point response. Acceptance of the manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of major revision only and acceptance or rejection of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

Given the related study on BioRxiv, it would be good if you could submit your revised manuscript as soon as possible. Would 1 or may be 2 months be sufficient for the revisions? Please let me know how much time you think you will need.

You can either publish the study as a short report or as a full article. For short reports, the revised manuscript should not exceed 27,000 characters (including spaces but excluding materials & methods and references) and 5 main plus 5 expanded view figures. The results and discussion sections must further be combined, which will help to shorten the manuscript text by eliminating some redundancy that is inevitable when discussing the same experiments twice. For a normal article there are no length limitations, but it should have more than 5 main figures and the results and discussion sections must be separate. In both cases, the entire materials and methods must be included in the main manuscript file.

Regarding data quantification, please specify the number "n" for how many independent experiments were performed, the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values in the respective figure legends. This information must be provided in the figure legends. Please also include scale bars in all microscopy images.

IMPORTANT NOTE: we perform an initial quality control of all revised manuscripts before re-review. Your manuscript will FAIL this control and the handling will be DELAYED if the following APPLIES:
1) A data availability section providing access to data deposited in public databases is missing. If you have not deposited any data, please add a sentence to the data availability section that explains that.
2) Your manuscript contains statistics and error bars based on n=2. Please use scatter blots in these cases. No statistics should be calculated if n=2.

When submitting your revised manuscript, please carefully review the instructions that follow below. Failure to include requested items will delay the evaluation of your revision.

1) a .docx formatted version of the manuscript text (including legends for main figures, EV figures and tables). Please make sure that the changes are highlighted to be clearly visible.

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure).
See https://wol-prod-cdn.literatumonline.com/pb-assets/embo-site/EMBOPress_Figure_Guidelines_061115-1561436025777.pdf for more info on how to prepare your figures.

3) We replaced Supplementary Information with Expanded View (EV) Figures and Tables that are collapsible/expandable online. A maximum of 5 EV Figures can be typeset. EV Figures should be cited as 'Figure EV1, Figure EV2" etc... in the text and their respective legends should be included in the main text after the legends of regular figures.

- For the figures that you do NOT wish to display as Expanded View figures, they should be bundled together with their legends in a single PDF file called *Appendix*, which should start with a short Table of Content. Appendix figures should be referred to in the main text as: "Appendix Figure S1, Appendix Figure S2" etc. See detailed instructions regarding expanded view here:

- Additional Tables/Datasets should be labeled and referred to as Table EV1, Dataset EV1, etc. Legends have to be provided in a separate tab in case of .xls files. Alternatively, the legend can be supplied as a separate text file (README) and zipped together with the Table/Dataset file.

4) a .docx formatted letter INCLUDING the reviewers' reports and your detailed point-by-point responses to their comments. As part of the EMBO Press transparent editorial process, the point-by-point response is part of the Review Process File (RPF), which will be published alongside your paper.

5) a complete author checklist, which you can download from our author guidelines . Please insert information in the checklist that is also reflected in the manuscript. The completed author checklist will also be part of the RPF.

6) Please note that all corresponding authors are required to supply an ORCID ID for their name upon submission of a revised

manuscript (). Please find instructions on how to link your ORCID ID to your account in our manuscript tracking system in our Author guidelines

7) Before submitting your revision, primary datasets produced in this study need to be deposited in an appropriate public database (see https://www.embopress.org/page/journal/14693178/authorguide#datadeposition). Please remember to provide a reviewer password if the datasets are not yet public. The accession numbers and database should be listed in a formal "Data Availability" section placed after Materials & Method (see also https://www.embopress.org/page/journal/14693178/authorguide#datadeposition). Please note that the Data Availability Section is restricted to new primary data that are part of this study. * Note - All links should resolve to a page where the data can be accessed. *
If your study has not produced novel datasets, please mention this fact in the Data Availability Section.

8) We would also encourage you to include the source data for figure panels that show essential data. Numerical data should be provided as individual .xls or .csv files (including a tab describing the data). For blots or microscopy, uncropped images should be submitted (using a zip archive if multiple images need to be supplied for one panel). Additional information on source data and instruction on how to label the files are available at .

9) Our journal also encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at https://www.embopress.org/page/journal/14693178/authorguide#referencesformat

We would also welcome the submission of cover suggestions, or motifs to be used by our Graphics Illustrator in designing a cover.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File (RPF) to accompany accepted manuscripts. This File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

You are able to opt out of this by letting the editorial office know (emboreports@embo.org). If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

I look forward to seeing a revised version of your manuscript when it is ready. Please let me know if you have questions or comments regarding the revision.

Kind regards,
Esther

Esther Schnapp, PhD
Senior Editor
EMBO reports

Revision 0

Review #1

1. How much time do you estimate the authors will need to complete the suggested revisions:

Estimated time to Complete Revisions (Required)

(Decision Recommendation)

Between 3 and 6 months

2. Evidence, reproducibility and clarity:

Evidence, reproducibility and clarity (Required)

In this manuscript, Anna Manjón et al. describe an unexpected outcome of CRISPR-Cas9 that results in the activation of the gene targeted by the sgRNA. This occurs via the integration of the lentivirus-based sgRNA vector into the endogenous target locus. The authors use the acquisition of taxol resistance as a readout for ABCB1 activation upon induction of a DSB in different regions of the gene using CRISPR-Cas9. While the experimental evidence for target gene activation is solid, data supporting the proposed model that sgRNA vector integration is the driver of such activation is scarce. The additional experiments (realistic, in my view) that are necessary to support the authors' claims are detailed below.

Data and the methods are presented in such a way that they can be reproduced, the experiments are adequately replicated and statistical analysis is adequate.

Additional experiments/analyses:

- The authors should provide additional information about the promoter region of the ABCB1 locus targeted by Cas9. Namely, CpG content, repressive histone modifications and the existence of binding sites for transcriptional repressors. The CpG methylation status and the presence of repressive chromatin marks are relevant as DSB-repair could lead to the erasure of the DNA methylation or histone modifications, causing gene activation independently of the vector integration. In addition, a putative transcriptional repressor could be physically displaced by Cas9 or have its binding site disrupted by mutagenesis upon repair of the DSB. In both scenarios, the outcome would be gene activation without the need for vector integration.
- Does Cas9 targeting the ABCB1 promoter lead to gene activation (i.e. taxol resistant colonies) if sgRNAs are provided using lipofectamine of electroporation?
- Is the DSB necessary? If the experiment is performed using a deadCas9 (+ the LV sgRNA vector) are there any taxol-resistant colonies?
- The authors suggest that 3 out of 1000 cells display ABCB1 gene activation and become taxol resistant. Do all resistant cells have the vector integrated into the promoter region? Can the authors determine the exact frequency of integration?

3. Significance:

Significance (Required)

This study is a significant contribution to the field. It will be relevant for a vast community of scientists that use CRISPR tools for a number of different purposes, such as genomic screens, genome editing, investigate DNA damage repair mechanisms, etc.

In my own field (the coupling between DNA damage and transcription) being aware of this unexpected outcome of CRISPR-Cas9 is definitely relevant, as many studies are now being conducted to assess the role of transcription during DSB repair, using Cas9 as the DNA damaging tool.

Referee Cross-commenting

Indeed, methods are described in a very succinct manner. I found it sufficient for reproducibility (especially because there are no new methods/protocols), but I do understand the concern raised by the other Reviewer. I am glad to withdraw my comments about the methods, if needed be.
I find all issues raised by the other Reviewer relevant and fair.

Review #2

1. How much time do you estimate the authors will need to complete the suggested revisions:

Estimated time to Complete Revisions (Required)

(Decision Recommendation)

Between 1 and 3 months

2. Evidence, reproducibility and clarity:

Evidence, reproducibility and clarity (Required)

In their paper "Unexpected gene activation following CRISPR-Cas9-mediated genome editing" Manjón et al. described a phenomenon in which Cas9 induced dabble strand brakes (DSBs) leading to integration of constitutive active promoter from the viral vector and activation of the nearby gene. It is known that viral cDNA can be integrated into DSBs, as this is the basis for the integration-defective lentiviral vector (IDLV) approach. Furthermore, it is known that viral integration may affect nearby gene expression due to viral enhancer/promoters, disruption of DNA regulatory elements, or epigenetic changes. In This paper, Manjón et al. highlight the possibility of aberrant gene activation by integrating truncated cDNA containing the promoter of the transgene. While, as far as I know, this specific type of "off-target" activity was never bean

described, however, the amount of data and its generality is limited.

Major comments

1. To understand if this phenomenon results from the specific genetics of the ABCB1 gene or more general phenomena, it is important to test if this can also happen with other genes.

2. The authors reporting this phenomenon only in one cell line. It is important to see whether these results can be reproduced in other cell lines.

3. It will be helpful if the authors can estimate the frequency of this phenomenon, so others can design control experiments to ensure that this type of off-target does not affect their results.

Minor comments

1. there is no figure legend for Figure 2F, and it is not clear.

2. Some statements in the introduction need references (e.g., "Combining imaging and high throughput technologies with DSB-induced Cas9 systems allows one to examine processes such as transcription, chromatin dynamics, and DNA replication.")

3. "Viral vectors include adeno-associated viruses (AAVs) and lentiviruses (LVs)". Also Gama retroviruses, and Adenoviral vectors are used for CRISPR.

4. "Chemical delivery methods comprise a DNA or mRNA form of the sgRNA that can be transfected into the host by lipofectamine reagents." Different liposome-based and non-leptosomic regents can be used for plasmid delivery. Lipofectamine is one transfection reagent provide by Invitrogen.

5. "In these latter strategies, the transfection efficiency can be lower, but they are a safer alternative, as random viral integrations do not occur." This statement is incorrect; plasmid DNA can randomly break/ linearized and integrate into the genome.

6. It will be nice if the authors will add few details regarding the RPE-1 cells. What kind of cells (cancer/normal, eye/kidney...)?

7. Using the lentiviral system lentiGuide-Puro from the Zhang Lab, we cloned different sgRNAs targeting different non-coding regions across the ABCB1 locus to induce a DSB (Fig 1A). add a reference to the lentiGuide-Puro (Sanjana 2018).

8. Materials and Methods are not detailed, hence hardly reproducible.

3. Significance:

Significance (Required)

This manuscript may be relevant both for LV base gene therapy and for academic studies. This off-target may affect oncogenes if generalized, increasing the risk for tumorgenicity (as described previously for gamma-retroviruses). This off-target may also affect the results of genetically engined models. Hence, it may lead to

misinterpretation of the observation and poor reproducibility.

Referee Cross-commenting

I find the other Reviewer's comments important and fair. the only point I disagree is regarding the method section, which I find short and not explicit.

1. General Statements [optional]

In this manuscript, we describe a mechanism via which genome-editing by CRISPR-Cas9 can lead to gene activation. This was a serendipitous finding that we uncovered when trying to document epigenetic changes that result from CRISPR-Cas9-mediated genome editing in the *ABCB1* locus. We found that the lentiviral-based sgRNA vectors integrate into the endogenous genomic target location at a low frequency, leading to undesired activation of the target gene. We were able to screen for this event as cells with *ABCB1* gene activation acquire resistance to Taxol. We believe that the scientific community should be alerted to this previously unreported side effect of CRISPR-Cas9-mediated genome editing.

2. Point-by-point description of the revisions

Reviewer #2 (Evidence, reproducibility and clarity (Required)):

In this manuscript, Anna Manjón et al. describe an unexpected outcome of CRISPR-Cas9 that results in the activation of the gene targeted by the sgRNA. This occurs via the integration of the lentivirus-based sgRNA vector into the endogenous target locus. The authors use the acquisition of taxol resistance as a readout for ABCB1 activation upon induction of a DSB in different regions of the gene using CRISPR-Cas9. While the experimental evidence for target gene activation is solid, data supporting the proposed model that sgRNA vector integration is the driver of such activation is scarce. The additional experiments (realistic, in my view) that are necessary to support the authors' claims are detailed below.

Data and the methods are presented in such a way that they can be reproduced, the experiments are adequately replicated and statistical analysis is adequate.

Additional experiments/analyses:

- The authors should provide additional information about the promoter region of the ABCB1 locus targeted by Cas9. Namely, CpG content, repressive histone modifications and the existence of binding sites for transcriptional repressors. The CpG methylation status and the

presence of repressive chromatin marks are relevant as DSB-repair could lead to the erasure of the DNA methylation or histone modifications, causing gene activation independently of the vector integration. In addition, a putative transcriptional repressor could be physically displaced by Cas9 or have its binding site disrupted by mutagenesis upon repair of the DSB. In both scenarios, the outcome would be gene activation without the need for vector integration.
We have now generated a set of H3k9me3 ChIP-seq experiments in RPE-1 cells (Parental and Taxol Resistant) to characterize the ABCB1 genomic region. We have also performed ChIP-qPCRs for the same repressive mark to further characterize the promoter region. We have also included a DNA Methylation array to assess the methylated CpG islands in the ABCB1 promoter. Here we also compared the parental cell line with the seven taxol resistant clones.
In these experiments we observed that some of the clones lose repressive chromatin marks but others maintain them. This shows that mere removal of chromatin modifications cannot fully explain ABCB1 gene activation. We have now added these experiments in the manuscript (see **Fig. 3**). In addition, we have used dCas9 to rule out the possibility that Cas9-dependent displacement of a transcriptional repressor is causing taxol-resistance (**Sup. Fig.1E**).

- Does Cas9 targeting the ABCB1 promoter lead to gene activation (i.e. taxol resistant colonies) if sgRNAs are provided using lipofectamine of electroporation?
We have generated synthetic crRNAs that contain the same targeting sequence as the Lenti-guide sgRNAs targeting the ABCB1 promoter. These synthetic guides were co-transfected with tracrRNA using lipofectamine reagents. We observed that taxol resistant colonies are still generated but with a much lower frequency than when using a lentiguide system. In the clones obtained with synthetic sgRNAs, we find that ABCB1 can also upregulated after a DSB due to genetic re-arrangements (promoter fusions). We are currently trying to collect the evidence that chromatin modifications induced during DSB repair could also play a role. However, as pointed out before, when using non-viral systems, the frequency of taxol-resistant colonies is much lower, as well as the expression levels of ABCB1 that are achieved.

Also with plasmid transfection you have this problem
In a preprint recently published in Biorxiv by Geng K et al 2021 (https://doi.org/10.1101/2021.07.01.450727) similar CRISPR off-target effects have been described. In this case Geng K et al generate the DSB by plasmid transfection and they also observe promoter integrations in the endogenous genome.

- Is the DSB necessary? If the experiment is performed using a deadCas9 (+ the LV sgRNA vector) are there any taxol-resistant colonies? We have now generated RPE-1 with a dCas9 construct. We have performed parallel infections with the LV sgRNA vectors in both RPE-1 Cas9 and dCas9 followed by Colony Formation Assays to assess the number of Taxol resistant cells. We observe that only when using a functional Cas9 Taxol resistant cells are generated, suggesting that the DSB is necessary to induce ABCB1 gene activation (see **Sup. Fig. 1C-E**).

- The authors suggest that 3 out of 1000 cells display ABCB1 gene activation and become taxol resistant. Do all resistant cells have the vector integrated into the promoter region? Can the

authors determine the exact frequency of integration? We know that not all Taxol resistant colonies have the vector integration in the promoter; DSB-dependent genomic rearrangements (such as gene fusions) can also lead to ABCB1 gene activation. However, we have found that these colonies grow slower under Taxol pressure compared to the vector integrated ones (see example CFA **Sup. Fig. 1A**). We have now incorporated data of these smaller Taxol-resistant colonies (**Sup. Fig. 3**). We have determined that these new clones derived from small colonies do not have the lentiguide puro integration vector integration and also are less resistant to Taxol, which correlates with lower mRNA levels of ABCB1 (**Sup. Fig. 3**). Therefore, in order to calculate the frequency of vector integration we have calculated the frequency of big Taxol resistant colonies obtained from the colony formation assays in Fig. 1A and Sup. Fig. 1A. In **Fig. 2G** we show the estimated average frequency (in %) of vector integration in the different sgRNAs targeting ABCB1 promoter.

Reviewer #2 (Significance (Required)):

This study is a significant contribution to the field. It will be relevant for a vast community of scientists that use CRISPR tools for a number of different purposes, such as genomic screens, genome editing, investigate DNA damage repair mechanisms, etc.

In my own field (the coupling between DNA damage and transcription) being aware of this unexpected outcome of CRISPR-Cas9 is definitely relevant, as many studies are now being conducted to assess the role of transcription during DSB repair, using Cas9 as the DNA damaging tool.

****Referee Cross-commenting****

Indeed, methods are described in a very succinct manner. I found it sufficient for reproducibility (especially because there are no new methods/protocols), but I do understand the concern raised by the other Reviewer. I am glad to withdraw my comments about the methods, if needed be.
I find all issues raised by the other Reviewer relevant and fair.

Reviewer #3 (Evidence, reproducibility and clarity (Required)):

In their paper "Unexpected gene activation following CRISPR-Cas9-mediated genome editing" Manjón et al. described a phenomenon in which Cas9 induced dabble strand brakes (DSBs) leading to integration of constitutive active promoter from the viral vector and activation of the nearby gene. It is known that viral cDNA can be integrated into DSBs, as this is the basis for the integration-defective lentiviral vector (IDLV) approach. Furthermore, it is known that viral integration may affect nearby gene expression due to viral enhancer/promoters, disruption of DNA regulatory elements, or epigenetic changes. In This paper, Manjón et al. highlight the possibility of aberrant gene activation by integrating truncated cDNA containing the promoter of

the transgene. While, as far as I know, this specific type of "off-target" activity was never bean described, however, the amount of data and its generality is limited.
We have added additional data supporting our findings, including data providing support for generality; we now show that the same phenomenon can be observed in different loci, and in different cell types.

Major comments

1. To understand if this phenomenon results from the specific genetics of the ABCB1 gene or more general phenomena, it is important to test if this can also happen with other genes.
We have performed experiments where we target the ABCG2 gene. ABCG2 is another drug efflux pump gene similar to ABCB1, whose upregulation leads to better efflux of Hoechst (CW Scharenberg, 2002). Hoechst is a fluorescent dye which allows for Flow Cytometry quantification of cells with high-low levels of this compound. We have validated that overexpression of ABCG2 via dCas9-Suntag-VP64 (CRISPRa) leads to lower levels of Hoechst in cells in RPE-1 cells (**Sup. Fig. 2B-D**). Next, we induced a DSB in the promoter of ABCG2 with a lentiviral system and sorted the low Hoechst population. Indeed, when performing PCRs over the DSB site in this polyclonal population we could observe higher bands appearing, which contained the EEF1A promoter from the LentiGuide vector (**Sup. Fig. 2E-H**). This indicates that vector integration can also happen in other genes which will lead to gene upregulation.
Also, Geng K et al 2021 (biorxiv) induce a DSB in a transfer RNA gene and they observe similar integrations in this locus.

2. The authors reporting this phenomenon only in one cell line. It is important to see whether these results can be reproduced in other cell lines.
We have also generated data targeting the ABCB1 promoter in Human Mammary Epithelial Cells (HMEC). Here we also observe Taxol resistant colonies only when we target the ABCB1 promoter with the lentiviral system (**Sup. Fig. 1B**). We have also picked these Taxol resistant colonies from sgRNA #12 and performed PCRs to assess vector integration. Indeed, as in RPE-1 cells, here we also observe vector integration in the DSB region (**Sup. Fig. 2A**).

3. It will be helpful if the authors can estimate the frequency of this phenomenon, so others can design control experiments to ensure that this type of off-target does not affect their results.
We have determined the frequency of integration per sgRNA in **Fig. 2G.**
Minor comments

1. there is no figure legend for Figure 2F, and it is not clear.
Apologies for the unclarity. We have now corrected the figure legend.

2. Some statements in the introduction need references (e.g., "Combining imaging and high throughput technologies with DSB-induced Cas9 systems allows one to examine processes such as transcription, chromatin dynamics, and DNA replication.")
We have now added references to this statement.

3. "Viral vectors include adeno-associated viruses (AAVs) and lentiviruses (LVs)". Also Gama retroviruses, and Adenoviral vectors are used for CRISPR.
We have now included Retroviral and Adenoviral vectors to this sentence.

4. "Chemical delivery methods comprise a DNA or mRNA form of the sgRNA that can be transfected into the host by lipofectamine reagents." Different liposome-based and non-leptosomic regents can be used for plasmid delivery. Lipofectamine is one transfection reagent provide by Invitrogen.
We have changed the word “lipofectamine” by liposome-based reagents.

5. "In these latter strategies, the transfection efficiency can be lower, but they are a safer alternative, as random viral integrations do not occur." This statement is incorrect; plasmid DNA can randomly break/ linearized and integrate into the genome.
We have now changed the statement for “*With RNA delivery methods, the transfection efficiency can be lower, but they are a safer alternative, as random viral integrations do not occur.*”

6. It will be nice if the authors will add few details regarding the RPE-1 cells. What kind of cells (cancer/normal, eye/kidney...)?
RPE-1 cells are Human Retinal Pigment Epithelial cells which have been immortalized by incorporation of the telomerase reverse transcriptase (hTERT). We have now included details of this cell line as well as Human Mammary Epithelial Cells (HMEC) in the method section.

7. Using the lentiviral system lentiGuide-Puro from the Zhang Lab, we cloned different sgRNAs targeting different non-coding regions across the ABCB1 locus to induce a DSB (Fig 1A). add a reference to the lentiGuide-Puro (Sanjana 2018).
Thank you. We have now added this reference.

8. Materials and Methods are not detailed, hence hardly reproducible.
Apologies for this. We have now clarified the materials and methods section and added new sections to better explain all the experiments performed.

Reviewer #3 (Significance (Required)):

This manuscript may be relevant both for LV base gene therapy and for academic studies. This off-target may affect oncogenes if generalized, increasing the risk for tumorgenicity (as described previously for gamma-retroviruses). This off-target may also affect the results of genetically engined models. Hence, it may lead to misinterpretation of the observation and poor reproducibility.

****Referee Cross-commenting****

I find the other Reviewer's comments important and fair. the only point I disagree is regarding the method section, which I find short and not explicit.

Dear Prof. Medema,

Thank you for the submission of your revised manuscript. We have now received the enclosed reports from the referees and I am happy to say that both support its publication now. We can therefore in principle accept it.

Only a few editorial requests still need to be addressed:

- Please correct the reference style. Up to 10 authors should be listed before "et al".

- Please remove the comment "data not shown" on page 4, as per journal policy.

- Fig 2G is called out after 3C. Fig 3B is callout out before 3A. There is a callout to Fig 4A+B, but Fig 4 has no panels. The Appendix figure panels are not called out. Please correct.

- Table 1 needs a title or legend that can be added to the first tab of the excel file.

- I attach to this email a related manuscript file with comments by our data editors. Please address all comments in the final manuscript.

- Please correct the before-last sentence in the abstract to present tense.

EMBO press papers are accompanied online by A) a short (1-2 sentences) summary of the findings and their significance, B) 2-3 bullet points highlighting key results and C) a synopsis image that is exactly 550 pixels wide and 200-600 pixels high (the height is variable). You can either show a model or key data in the synopsis image. Please note that text needs to be readable at the final size. Please send us this information along with the revised manuscript.

I look forward to seeing a new revised version of your manuscript as soon as possible. Please use this link to submit your revision: https://embor.msubmit.net/cgi-bin/main.plex

Best regards,
Esther

Esther Schnapp, PhD
Senior Editor
EMBO reports

Referee #1:

In this revised manuscript, the authors addressed all my concerns. In my opinion, the manuscript is now ready for publication.

Referee #2:

Manjón et al. described a phenomenon in which Cas9 induced dabble strand brakes (DSBs) leading to integration of constitutive active promoter from the viral vector and activation of the nearby gene. Although the frequency of these event are very low, and observed to other DSB events, this highlight an important off-target result.
The authors answered the concerns raised by this reviewer.

The authors have addressed all minor editorial requests.

Prof. Rene Medema
Oncode Institute, the Netherlands Cancer Institute
Department of Cell Biology
Plesmanlaan 121
Amsterdam 1066 CX
Netherlands

Dear Prof. Medema,

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports. Thank you for your contribution to our journal.

Given that you as corresponding author are in the Netherlands, you are probably eligible for publication of your article in the open access format in a way that is free of charge for the authors. Please contact either the administration at your institution or our publishers at Wiley (emboreports@wiley.com) for further questions. The general information of Wiley's agreements with the Netherlands can be found here:
https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/affiliation-policies-payments/vsnu-agreement.html

At the end of this email I include important information about how to proceed. Please ensure that you take the time to read the information and complete and return the necessary forms to allow us to publish your manuscript as quickly as possible.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File to accompany accepted manuscripts. As you are aware, this File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

If you do NOT want this File to be published, please inform the editorial office within 2 days, if you have not done so already, otherwise the File will be published by default [contact: emboreports@embo.org]. If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

Thank you again for your contribution to EMBO reports and congratulations on a successful publication. Please consider us again in the future for your most exciting work.

Best wishes,
Esther

Esther Schnapp, PhD
Senior Editor
EMBO reports

****

THINGS TO DO NOW:

You will receive proofs by e-mail approximately 2-3 weeks after all relevant files have been sent to our Production Office; you should return your corrections within 2 days of receiving the proofs.

Please inform us if there is likely to be any difficulty in reaching you at the above address at that time. Failure to meet our deadlines may result in a delay of publication, or publication without your corrections.

All further communications concerning your paper should quote reference number EMBOR-2021-53902V3 and be addressed to emboreports@wiley.com.

Should you be planning a Press Release on your article, please get in contact with emboreports@wiley.com as early as possible, in order to coordinate publication and release dates.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Rene H Medema
Journal Submitted to: EMBO Reports
Manuscript Number: EMBOR-2021-53902V2

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	GraphPad StatMate was used to choose sample size for at least 80% statistical power and a 95% confidence interval.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	No samples were excluded.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	No randomization was required as in every experiment, all samples were assessed simultaneously.
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	All samples of one experiment were processed simultaneously (with same batch solution, reagents, antibodies etc). The aquisition (microscope and flow cytometry) was performed using the same settings for all samples in one experiment.
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Statistical analysis was not performed in this study
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	NA
Is there an estimate of variation within each group of data?	NA

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repor

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumo

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is the variance similar between the groups that are being statistically compared?	NA

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	See Materials and Methods of the manuscript for information about the used antibodies
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	RPE and HMEC cells were originally from ATCC. Cells were regulary tested for mycoplasm. Information about the generation of Cas9 cells is reported in Materials and Methods.

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	d. In Materials and Methods the GEO accession number of ChIP-sequencing experiments and TLA can be found. GEO accession code is GSE185725.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	See above
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	No