nature biomedical engineering

PEER-REVIEW INFORMATION

Large DNA deletions occur during DNA repair at 20-fold lower frequency for base editors and prime editors than for Cas9 nucleases

Corresponding author: Sangsu Bae

Editorial note

This document includes relevant written communications between the manuscript's corresponding author and the editor and reviewers of the manuscript during peer review. It includes decision letters relaying any editorial points and peer-review reports, and the authors' replies to these (under 'Rebuttal' headings). The editorial decisions are signed by the manuscript's handling editor, yet the editorial team and ultimately the journal's Chief Editor share responsibility for all decisions.

Any relevant documents attached to the decision letters are referred to as **Appendix #**, and can be found appended to this document. Any information deemed confidential has been redacted or removed. Earlier versions of the manuscript are not published, yet the originally submitted version may be available as a preprint. Because of editorial edits and changes during peer review, the published title of the paper and the title mentioned in below correspondence may differ.

Correspondence

Tue 06 Feb 2024
Decision on Article nBME-24-0092-T

Dear Prof Bae,

Thank you again for submitting to *Nature Biomedical Engineering* your manuscript, "Detailed mechanisms for unintended large DNA deletions with CRISPR, base editors, and prime editors". The manuscript has been seen by three experts, whose reports you will find at the end of this message. You will see that the reviewers appreciate the work, and that they raise a number of technical criticisms that we hope you will be able to address. In particular, we would expect that a revised version of the manuscript provides:

* Additional mechanistic evidence, in particular that by which the PE3 system leads to a higher incidence of large deletions, as per the comments and suggestions from Reviewers #1 and #2.

* Further assessment of the degree of reliability of the alignment algorithm, particularly through comparisons with other alignment algorithms and by providing evidence from additional endogenous sites, as recommended by Reviewers #1 and #3.

* A carefully nuanced discussion of the extent and clinical relevance of the large deletions in the base editing and prime editing systems, particularly when compared to those by Cas9-based editors producing double-strand breaks.

* Comprehensive methodological reporting.

When you are ready to resubmit your manuscript, please upload the revised files, a point-by-point rebuttal to the comments from all reviewers, the reporting summary, and a cover letter that explains the main improvements included in the revision and responds to any points highlighted in this decision.

Please follow the following recommendations:

* Clearly highlight any amendments to the text and figures to help the reviewers and editors find and understand the changes (yet keep in mind that excessive marking can hinder readability).

* If you and your co-authors disagree with a criticism, provide the arguments to the reviewer (optionally, indicate the relevant points in the cover letter).

* If a criticism or suggestion is not addressed, please indicate so in the rebuttal to the reviewer comments and explain the reason(s).

* Consider including responses to any criticisms raised by more than one reviewer at the beginning of the rebuttal, in a section addressed to all reviewers.

* The rebuttal should include the reviewer comments in point-by-point format (please note that we provide all reviewers will the reports as they appear at the end of this message).

* Provide the rebuttal to the reviewer comments and the cover letter as separate files.

We hope that you will be able to resubmit the manuscript within 12 weeks from the receipt of this message. If this is the case, you will be protected against potential scooping. Otherwise, we will be happy to consider a revised manuscript as long as the significance of the work is not compromised by work published elsewhere or accepted for publication at *Nature Biomedical Engineering*.

We hope that you will find the referee reports helpful when revising the work, which we look forward to receive. Please do not hesitate to contact me should you have any questions.

Best wishes,

Pep

__
Pep Pàmies
Chief Editor, *Nature Biomedical Engineering*

_____
Reviewer #1 (Report for the authors (Required)):

This manuscript by Gue-ho Hwang et al., entitled 'Detailed Mechanisms for Unintended Large DNA Deletions with CRISPR, Base Editors, and Prime Editors,' utilizes an optimized long-range amplicon sequencing system and develops a k-mer sequence-alignment algorithm to simultaneously detect small DNA alteration events and large DNA deletions. The authors determined the occurrence of large deletions in different cell lines and verified that TMEJ is the major process producing large deletions. Furthermore, using a similar method, they discovered large deletions in systems utilizing base editors and prime editors and proposed hypotheses to explain how small indels or large deletions occur in base editing systems.

Verifying the function of TMEJ represents an important discovery in this report; however, some observations presented in this manuscript have been similarly reported elsewhere. For example, the phenomenon of large deletions induced by CRISPR/Cas9 was already proposed by Fatwa Adikusuma et al. (PMID: 30089922), and similar hypotheses regarding indels in the base editing (BE) system have been mentioned elsewhere (PMID: 32690971; PMID: 27096365). Additionally, the occurrence of indels in H840A-nCas9, which is used in PE2 or PE3, has been previously proposed (PMID: 36997524). In addition, the nCas9 could induce indels mutation with two sgRNAs is also already confirmed (PMID: 23992846), thus two sgRNA might increasing the likelihood of large indels (>100bp, as mentioned in the manuscript) during the PE3 editing process.

Without a deeper clarification of the mechanism, retelling similar cases does not add novelty. In conclusion, I believe this article now does not meet the publication standards of NBE, and I suggest that more detailed

mechanistic experiments should be conducted to enhance the value of the article.

1. The term 'detailed mechanisms' has been proposed in the article title to describe the workings of 'CRISPR, base editors, and prime editors.' However, upon review, I found that the detailed mechanism of large deletions was only discussed in the context of CRISPR. For the base editor section, the mechanism of insertions and deletions (indels) was presented merely as a hypothesis by the author. Moreover, the 'detailed mechanism' behind the occurrence of large indels in prime editing has not been clearly elucidated in this manuscript. Moreover, POLQ significantly contributes to the repair of double-strand breaks via the theta-mediated end joining (TMEJ) pathway, however, might not be key regulator in the gene editing in base editors or primer editors, for these system relies on nickase Cas9. Therefore, the term 'detailed mechanism' might only be applicable to CRISPR, but not to base editing or prime editing, as presented in this manuscript. If the authors still wish to use this term, further experiments are needed to more thoroughly clarify the mechanisms by which large indels occur in both base editing and prime editing.

2. In the result section “Optimizing a long-range amplicon sequencing method to determine small indels and large deletions with a high accuracy”, The author highlights the high accuracy of the Illumina platform, which justifies the initial use of large fragment PCR followed by the development of specific algorithms for indel matching base on NGS platform. However, when it comes to the methodology for constructing long fragments, the approach lacks significant innovation. Conducting PCR on a 10kb genomic length is relatively straightforward. For example, in the full-length mitochondrial genome off-target detection process (PMID: 32641830), after long range PCR amplication (2 PCR, total nearly 16kb), by employing ultrasonication for random fragmentation followed by library construction and sequencing, it is feasible to assemble the sequences and identify indels or SNVs across the entire fragment. This technique is already well-established in the field. Consequently, I find that this method does not introduce notable innovation in the context of existing techniques.

From an algorithmic perspective, the author should highlight the advancements in accuracy and convenience offered by the newly established algorithm, rather than primarily emphasizing its "high speed and less memory" in the main text and Figure 1a. In the context of detecting indels within a 10kb DNA sequence, if indexes are exclusively created for 10Kb segments, utilizing BWA for alignment and subsequent analysis allows for rapid completion, without significantly taxing hardware memory. On the other hand, the claims of "high speed and less memory" made by the author require validation through comparative analysis with other algorithms.

Additionally, the reliability of this algorithm is fundamental to the entire article. Therefore, using limited endogenous sites (only two, in this manuscript) to validate the reliability of a method is insufficient. The author needs more evidence to demonstrate that this newly established algorithm can effectively search for large deletions.

In addition, the reliable of the k-mer_align tool should be compare with other software or algorithm. For instance, previous algorithms have been shown in literature not to produce significant large indels in the BE system (PMID: 32711379), yet the results in this article suggest the opposite. The algorithm may be a crucial factor contributing to these differing outcomes. If the author need to be more invincible, I believe further parallel validation of the algorithm is essential. If this point are not proved well, then in every subsequent result, I recommend adding steps to validate with other algorithms (possibly in the supplementary material) to confirm the reliability of the findings.

Furthermore, the author could synthetically create or establish more complex library sequences, including both small indels and large deletions, to elucidate this issue in depth, making your results more credible.

3. The preference of PCR enzymes is important, but I noticed that the gradient of preference used by the author is: 120, 140, 160, 180, 200, 220, 240, 260, 280bp. Given that the author's goal involves large fragments of 15kb and the focus is on large deletions (> 100bp), this gradient does not significantly contribute evidence towards enzyme preference for this article. It is necessary to include fragments with greater differences, such as 100bp, 500bp, 1000bp, and 10000bp, to conduct relevant experiments and further substantiate the author's point. Moreover, the extension time should be tested when performing PCR on longer sequences since it is possible that the extension time could be a key factor in limiting the length of the PCR product and may contribute to bias in the length of the PCR product.

Additionally, for the author's experiment, the ability to successfully perform PCR on 10kb genomic fragments

is also very important. Evidence (gel running) of successful long range PCR with these five enzymes needs to be provided (which can be included in the supplementary material).

4. In lines 167-169, it is observed that the CRISPRi library used by the authors specifically targeted 794 genes referenced in Repair-seq. However, in Figure 3B, only the genes associated with End protection and End resection are presented. This raises a question: what about the other genes? It is difficult to believe that only the Z-scores of End protection and End resection have undergone changes. What if there are additional genes that play a crucial role in DSB processes? To address this concern, it would be valuable for the authors to include a figure displaying the Z-scores of other significantly altered genes not belonging to the End protection and End resection categories. Furthermore, it is important for the authors to engage in a discussion regarding these findings to fully explore the potential implications.

5. The selection of candidate genes depicted in Figure 3e for investigating the major pathway linked to large deletions raises some questions. For instance, it is notable that ERCC4, a gene within the SSA gene group, was not included as a candidate gene. Furthermore, there is a lack of gene knockout experiments conducted within the HR gene group. It is crucial to provide a clear explanation for the reasons behind these selection and experimental decisions. If you claim that another pathway is not the major pathway for large deletions to occur in gene editing, it is necessary to provide direct evidence to support this exclusion.

6. Figure 3e demonstrates that gene silencing of PQLQ could significantly decrease the occurrence of large deletions, implying that TMEJ is the primary pathway involved in large deletions. To further support this conclusion, it is essential to confirm whether overexpression of PQLQ could increase the incidence of large deletions in CRISPR gene editing. This experiment would add substantial strength to the author's argument and help to further establish the role of TMEJ in the mechanism of large deletions.

7. In lines 185-188, the authors stated, "Using a previously defined classification scheme, we categorized the genes into those associated with NHEJ (LIG4, XRCC4, XRCC5, and XRCC6), end resection (MRE11A, NBN, and RAD50), TMEJ (POLQ), SSA (ERCC4 and RAD52), and HR (BRCA1, BRCA2, RAD51AP1, RAD51B, RAD51C, and RAD51D)." From this perspective, individually knocking out each of these genes might be a more suitable approach than CRISPRi screening.

8. In line 232,the author mentioned “Collectively, these data support evaluation of M4344 or drugs that reduce end resection or TMEJ pathways as co-effectors with CRISPR-Cas9-mediated genetic engineering in therapeutic applications to mitigate unintended large deletions”. It’s a good point that that M4344 or drugs targeting the end resection or TMEJ pathways could potentially be used alongside CRISPR-Cas9-mediated genetic engineering to mitigate unintended large deletions in therapeutic applications. However, it should be noted that the primary function of M4344 is to pharmacologically inhibit ATR activity, thereby suppressing the entire end resection process, including TMEJ, SSA, and HR. In this study, the author did not employ any specific strategy to individually inhibit the TMEJ pathway, such as POLQ inhibitor (ART558) in primary T cells.

9. In the section titled "Confirmation by knockout that end resection and TMEJ processes generate large deletions", I observed that the author solely focuses on the decreased occurrence of large deletions during the gene editing process. However, it is crucial to also examine the on-target efficiency of the M4344+ or M4344- groups in the tested samples. Was there any impact of M4344 on the normal editing efficiency of the target genes? If the on-target effect was indeed reduced, it would consequently lead to a decrease in off-target effects as well. The relationship between on-target efficiency and the occurrence of large deletions is a crucial factor to consider when determining the suitability of an assistant drug in gene therapy. To provide a comprehensive understanding, it is important for the author to include information about the on-target effects in the same figure. This would allow the readers to discern that the decrease in large deletions is not solely attributed to a decline in the editing efficiency of the gene editor.

10. In line 526, the author mentioned “ HeLa CRISPRi cells were generated by lentiviral integration (~3 to 5 MOI) using the dCas9-KRAB-blast plasmid (Addgene #89567), followed by single cell isolation.” However, this raises a question about the rationale behind the single cell isolation. It seems the procedure involves introducing a lentivirus library into the cells before isolating individual cells. If this is indeed the case, it implies the potential creation of over 795 unique cell lines, each corresponding to different single-guide RNAs (sgRNAs) for CRISPR interference (CRISPRi). Such a scenario suggests that 795 individual transfections of ribonucleoprotein complexes (RNPs) would be necessary for the screening experiment. Nevertheless, the text does not mention any details regarding the execution of these individual transfections,

leaving a gap in the explanation of the methodology.

Furthermore, the methodology for screening sgRNAs and confirming their association with the signaling pathway needs to be meticulously detailed in the methods section. It is equally important that a comprehensive description of these processes is presented in the results section to ensure clarity and understanding.

11. It is important to note that both base editors and prime editors utilize nCas9 as the underlying structure. In this manuscript, the base editors mentioned employ D10A nCas9 as the base architecture, while PE2 and PE3 use H840A nCas9. Considering this perspective, it is recommended to include a control group with nCas9 in Figure 4 and Figure 5, targeting different sites.

12. In Figure 4, it is essential to provide a detailed version and plasmid information of the cytosine base editors (CBE), cytosine-guanine base editors (CGBE), adenine base editors (ABE), and adenine-to-cytosine base editors (AYBE).

13. In Figure 4e, the author presents a hypothesis concerning the occurrence of small indels or large deletions. However, it is important to note that similar hypothesis or part of the hypothesis has already been proposed in previous reports (PMID: 32690971; PMID: 27096365). Thus, the hypothesis lacks novelty.

14. As mentioned earlier, there have been reports suggesting that base editing does not result in large deletions. It is commendable that the author has identified instances of large deletions occurring in base editors. However, the conflicting viewpoints need to be carefully addressed and discussed. To strengthen the argument, more compelling evidence and better discussion should be provided.

15. In Figure 5C, we observe that the efficiency of large deletions using PE3 is categorized into two distinct groups: one group exhibits an efficiency of approximately 3%, while the other demonstrates a significantly higher efficiency around 50%. This substantial variance warrants further discussion to understand the underlying causes.

Moreover, it is important to explore the potential relationship between these higher editing efficiency and the large deletion efforts of PE2/PE3.

16.The results of PE3+epeg, as shown in Figure 5d, exhibit slightly significant variability. To obtain more reliable data, increasing the sample size is recommended.

17. In Fig 5 and discussion section, further discussion is warranted regarding the underlying mechanism by which the PE3 system leads to a higher incidence of indels and large deletions. Previous studies have demonstrated that efficient targeting effects can be achieved through the combined action of two nCas9 enzymes (PMID: 23992846). As the PE3 system employs a dual sgRNA strategy, which shares similarities with the principle of gene knockout using nickase Cas9, it is crucial to include a control group in Figure 5. Specifically, adding nCas9 + sgRNA1 + sgRNA2 (where sgRNA1/2 corresponds to the sgRNAs utilized in PE3) might help ascertain whether the observed large indels are solely attributable to nCas9.

18. In Figure 5, it is evident that the authors have not provided mechanistic explanations or a hypothesis. This lack of information creates a disconnect from the "detailed mechanism" (in title) concerning prime editors. Therefore, it is crucial to design and incorporate additional experiments pertaining to the mechanism in Figure 5.

Reviewer #2 (Report for the authors (Required)):

The manuscript "Detailed mechanisms for unintended large DNA deletions with CRISPR, base editors, and prime editors," by Hwang, Lee, and colleagues sought to characterize the frequencies of and mechanisms by which large DNA deletion byproducts are generated during the course of genome editing experiments. To address the quantification problem, they developed a long-range next-generation sequencing (NGS) pipeline to investigate the genotypes that arise from Cas9 nuclease cutting, base editing, and prime editing. To investigate the mechanism by which these genotypes occur, they performed a CRISPRi screen and found that long deletion genotypes arise from the theta-mediated end-joining (TMEJ) pathway. Lastly, they

examined a variety of base editing architectures and prime editing systems to investigate the frequencies of long deletions in those systems.

The work addresses an important question in the application of therapeutic genome editing: "How can one quantify and reduce undesired genotypes that cannot be sequenced directly with traditional NGS techniques?" However, while this question is important, there are several aspects of this work that must be addressed before it is suitable for publication in Nature Biomedical Engineering. Addressing the following points would significantly strengthen the manuscript, and improve its relevance to a biomedical engineering audience:

Major points:
- (Lines 38–39): The maximum frequencies of the deletions that the authors found for biomedically relevant base editors (CBE: 0.8%, ABE: 1.42%) are quite small. Moreover, the frequencies of these deletions are low for prime editors as well, with a maximum large-deletion frequency of 1.2% for PE2 and a clear bimodal distribution for PE3, with the majority of edits having a large-deletion frequency under 5%. The abstract as written makes a misleading and false equivalence between Cas9-based and nCas9-based technologies as generating similar levels of such byproducts. In reality, the main finding, that BE and PE result in far lower frequencies of these large deletions, is less sensationalist but much more important and accurate. The authors should reword this sentence to make this distinction more clear or remove this sentence entirely from the abstract, as it is misleading and sensationalist.

- (232–234): While M4344 works for reducing TMEJ-mediated large deletions, inhibition of ATR kinase would also preclude the HDR mechanism and would therefore be unsuitable for approaches used in many applications of Cas9 nuclease, such as integration of a chimeric antigen receptor into T cells. The authors should make this caveat more clear. Alternatively, a more convincing argument could be made using an inhibitor of DNA polymerase theta to selectively inhibit TMEJ but not HDR.

- (249–253): Importantly, CBE(ΔUGI), CGBE, and AYBE are not relevant to clinical applications of base editing—they are explicitly avoided for clinical use because their mechanism yields byproducts including deletions (though still fewer than nucleases). A better way of structuring this section is to present the canonical CBE/ABE results (which show low deletion rates), and then describe the additional non-clinical constructs tested as part of a mechanistic analysis showing how these low-frequency deletions arise. Then, the authors can add that based on their analysis, ABE–UGI could be used in cases where there are TC motifs present in the base editing window to further drive down indels resulting from bystander cytidine deamination.

- (293–294): The authors should examine the exact deletion genotypes that arise from PE3. Do these deletions extend past the two nick sites? If so, this would suggest a mechanism that also involves resection mediated by ATR kinase and therefore treatment with M4344 should reduce the frequency of these deletion genotypes. If not, then the authors should note that the mechanism by which these genotypes occur is distinct from that discussed earlier in the paper, and perhaps propose a different mechanism that can explain the results observed.

- (452–496): Were read pairs deduplicated before processing? This step is crucial for quantitative analyses involving DNA shearing followed by a barcoding amplification step because PCR bias during the barcoding step can result in some sequences being duplicated and therefore overrepresented.

Minor points:
- (136, 486): In the main text, the authors defined the cutoff for large vs. small deletions at 100 bp. However, in the methods, the authors defined the cutoff at 50 bp. Can the authors please clarify this discrepancy?

- (318–322): New data should not be included in the discussion section. Instead, the authors can consider making a new results sub-section to discuss mischaracterization of cell lines. It would also strengthen the overall impact of the paper if the authors can make an analysis of the likelihood of such mischaracterization in current clinically relevant genome editing workflows.

- (452–496): The authors did not specify the k-mer length they used for analyses performed in this work. Additionally, the use of the letter k for both the sequence length in the k-mer hash table and the k-means clustering hyperparameter is confusing.

Reviewer #3 (Report for the authors (Required)):

This is a nice article in which the authors highlight the issue of long deletions (>100 bp) after gene editing. Although the long deletion phenomenon has been previously described (as authors cite), here the authors develop some improved experimental and analytic assays for long-range amplification, fragmented short-read sequencing or long-read sequencing, and edit sequence alignment, conduct a genetic screen identifying end resection as an important pathway required for long deletions, and demonstrate the risks of long deletions with certain base and prime editing approaches. This will be a useful reference for future investigators seeking to comprehensively characterize gene editing outcomes including long deletion events and is a cautionary tale that long deletions may indeed be present after various forms of gene editing if only one looks carefully.

Some specific comments:

1. Fig 1c shows minimal bias comparing a ~9kb and ~10kb product (~10% difference in length). Also Supp Fig 3 is a nice design, but only tests range of 120-280 bp. What about longer deletions, e.g. up to 9 kb deletions for a 10 kb product (~90% difference in length)? It would be interesting to see the performance of the polymerases around a range of larger deletions to better understand possible length biases of the procedures.

2. Regarding the performance of the alignment strategies, it could be interesting to simulate reads with various deletions and compare approaches.

3. In Fig 2, are the small insertions templated +1 position 17 reiterations or short homonucleotide insertions as have been described as a common Cas9 repair outcome (PMC5834694) or some other type of small insertion which might reflect a different repair mechanism? The anti-correlation might suggest competing DNA repair mechanisms.

4. Fig 5 and Supp Fig 9: it could be interesting to try PE3b strategy to see if it is more similar to PE2 or PE3 in terms of long deletions.

5. Could the authors add some discussion regarding to what degree including a UMI might increase the accuracy and minimize the bias of the long read quantification?

6. A few typos:
- Supp Fig1b "bluw" typo.
- Typo line 261-2 "In Figure 4c, each large deletion frequency was divided by the average large deletion frequency with."
- Fig S10 typo "Stard"
- For lines 321 and 322, is it meant "homozygous" and "heterozygous" rather than "homologous" and "heterologous"?

Tue 16 Jul 2024
Decision on Article NBME-24-0092A

Dear Prof Bae,

Thank you for your revised manuscript, "Unintended large DNA deletions occur through cellular DNA repair pathways during gene editing with CRISPR, base editors, and prime editors", which has been seen by the original reviewers. In their reports, which you will find at the end of this message, you will see that the reviewers acknowledge the improvements to the work and raise a few additional technical criticisms that we hope you will be able to address. In particular, we would expect that the next version of the manuscript provides a sufficiently nuanced discussion of the findings, also with regards to relevant literature, as per the recommendations of Reviewers #1 and #2. Although not a condition for the eventual publication of the study, additional mechanistic data and insight (as per the suggestions of Reviewer #1) would be appreciated.

As before, when you are ready to resubmit your manuscript, please upload the revised files, a point-by-point rebuttal to the comments from all reviewers, the reporting summary, and a cover letter that explains the main improvements included in the revision and responds to any points highlighted in this decision.

As a reminder, please follow the following recommendations:

* Clearly highlight any amendments to the text and figures to help the reviewers and editors find and understand the changes (yet keep in mind that excessive marking can hinder readability).

* If you and your co-authors disagree with a criticism, provide the arguments to the reviewer (optionally, indicate the relevant points in the cover letter).

* If a criticism or suggestion is not addressed, please indicate so in the rebuttal to the reviewer comments and explain the reason(s).

* Consider including responses to any criticisms raised by more than one reviewer at the beginning of the rebuttal, in a section addressed to all reviewers.

* The rebuttal should include the reviewer comments in point-by-point format (please note that we provide all reviewers will the reports as they appear at the end of this message).

* Provide the rebuttal to the reviewer comments and the cover letter as separate files.

We look forward to receive a further revised version of the work. Please do not hesitate to contact me should you have any questions.

Best wishes,

Pep

__
Pep Pàmies
Chief Editor, *Nature Biomedical Engineering*

_____
Reviewer #1 (Report for the authors (Required)):

Summary
The author has made several revisions to the article, each contributing to its enhanced clarity and depth. Firstly, they took steps to refine the title, ensuring it accurately encapsulates the content and purpose of the study, thereby facilitating easier comprehension for readers. Secondly, the author meticulously tailored detailed explanations to various algorithm optimizations, providing thorough insights that cater to different aspects of the research. Moreover, the author bolstered the content by introducing precise inhibitors and

including essential control groups, crucial for establishing comprehensive experimental frameworks. Additionally, they integrated mechanistic experiments of the PE editor, offering practical demonstrations that further substantiate the findings and methodologies discussed. They also diligently rectified corresponding descriptions throughout the article, improved accuracy and resolve some notable issues that might affect the scholarly integrity of the work. Overall, these revisions enhance the rigor and underscore the overall quality of the manuscript.

Major technical criticisms or questions
I still maintain the view that optimizing PCR protocols and algorithms does not constitute a significant innovation. Currently, in routine off-target detection of large segments, there are limited opportunities to evaluate large segment deletions across thousands of edits. The algorithm optimization proposed by the authors reduces computer resource usage, such as memory, from 90M in BWA-MEM to 17M in ExCas-Analyzer. However, this reduction in memory usage may not be particularly meaningful given the capabilities of current computer hardware resources (even personal laptops commonly have 16GB of memory). While I deeply appreciate the authors' efforts, I do not believe that optimizing PCR enzymes and algorithms warrants being highlighted as a standout feature in a prestigious journal like NBE.

Furthermore, I believe the article still presents opportunities for innovation. Previous studies have documented instances of large deletions in gene editing using CRISPR/Cas9 (PMID: 30089922) , deletions in base editing tools (PMID: 32690971; 27096365) and PE system (PMID: 36997524). The emphasis on large segment deletions in these editors lacks novelty. In contrast, while observations of such deletions with PE3 have been noted (specifically, nCas9's potential role in DNA fragment deletion, PMID: 23992846), there hasn't been a comprehensive demonstration that PE3 induces unintended large deletions. Therefore, highlighting the findings related to PE3 might be helpful to enhance the article's impact.

However, the mechanistic issues addressed by the authors continue to be a valuable component of this manuscript. A more comprehensive understanding of the mechanisms responsible for large segment deletions could assist in mitigating unintended editing effects and enhancing the safety of editing tools. This knowledge bears considerable importance for the clinical deployment of diverse editing technologies. However, I do feel some regret that despite the authors' inclusion of pertinent data, their investigation into these mechanisms lacks thoroughness, which detracts somewhat from the article's overall innovativeness.

Reviewer #2 (Report for the authors (Required)):

The manuscript "Unintended large DNA deletions occur through cellular DNA repair pathways during gene editing with CRISPR, base editors, and prime editors" by Hwang, Lee, and colleagues aimed to demonstrate that Cas9-based genome editing tools produce large DNA deletions that were previously difficult to characterize with short-read sequencing tools. To address this problem, they developed a custom next-generation sequencing (NGS) pipeline and used it to investigate the genotypes that arise from genome editing with Cas9 nucleases, base editors, and prime editors.

In their revision, the authors addressed several of the points that we raised previously. It was especially helpful that the authors clarified in lines 271-276 that the goal of using the specific BE architectures that they chose was to investigate the mechanism by which large deletions mediated by BEs arise. However, there are still a few important outstanding issues that must be corrected before the manuscript meets the standards of rigor and scholarship necessary for publication.

Major points:
- Line 2: The title remains quite misleading, because it implies that CRISPR nucleases, base editors, and prime editors all give similar levels of large DNA deletions. In reality there are large differences, as expected and previous established, between large DNA deletion frequencies from these three genome editing techniques. Figure 4d and Figure 5b show that both base editors and prime editors result in more than 20-fold lower average frequencies of large deletions than Cas9 nuclease! This important and basic finding is not reflected at all in their title and should also be reflected more accurately in their abstract. Other papers previously published with similarly misleading titles that are not consistent with the papers' data have been deservedly criticized after publication as sensationalist and lacking in scholarship. The authors should revise the title of the manuscript to more accurately reflect their findings. For example: "Different frequencies of unintended large DNA deletions from CRISPR nucleases, base editors, and prime editors", or "Unintended

large deletions occur 20-fold more frequently from CRISPR nucleases than base editors or prime editors".

- Line 40: Similarly, while I appreciate that the authors amended their abstract to say "fewer but significant" large deletions for base editors and prime editors, it is important that the authors acknowledge that they assayed base editors irrelevant to a biomedical engineering audience. Although I understand that "significant" is used here in the statistical sense, it would be more consistent with their data and less misleading and sensationalist if the authors instead changed this sentence to "Furthermore, the assay we developed was sensitive enough to also detect low levels of large deletions arising during base editing and prime editing."

- Line 234: It was interesting to see the use of M4344 and ART558 to reduce the frequency of large deletions generated by Cas9 nuclease. However, the authors did not actually demonstrate the feasibility of using either molecule to improve nuclease editing outcomes in any therapeutically relevant scenario. The authors could consider either adding such an experiment in (e.g., inserting a CAR into the TRAC locus of a T cell line with HDR), or alternative, they should remove "for therapeutic applications" from the section header. Many more data re required to establish true therapeutic application relevance.

Minor point:
- Lines 351-358: It is interesting that the large deletion frequency increased in the LIG4 KO line, given the canonical role of LIG4 in cNHEJ. The authors should consider commenting on this phenomenon, perhaps by looking at the distribution of genotypes observed instead of large deletions.

Reviewer #3 (Report for the authors (Required)):

The authors have addressed my concerns in this responsive revision.

Tue 03 Aug 2024
Decision on Article NBME-24-0092B

Dear Prof Bae,

Thank you for your revised manuscript, "Unintended large DNA deletions occur through cellular DNA repair pathways during gene editing with CRISPR, base editors, and prime editors". Having consulted with Reviewers #1 and #2 (whose comments you will find at the end of this message), I am pleased to write that we shall be happy to publish the manuscript in *Nature Biomedical Engineering*, provided that the points specified in the attached instructions file are addressed.

Please note that in the document I am providing an edited title and abstract that takes into account the comments by Reviewer #2.

When you are ready to submit the final version of your manuscript, please upload the files specified in the instructions file.

We encourage authors to take up transparent peer review. If you are eligible and opt in to transparent peer review, we will publish, as a single supplementary file, all the reviewer comments for all the versions of the manuscript, your rebuttal letters, and the editorial decision letters. **If you opt in to transparent peer review, in the attached file please tick the box 'I wish to participate in transparent peer review'; if you prefer not to, please tick 'I do NOT wish to participate in transparent peer review'.** In the interest of confidentiality, we allow redactions to the rebuttal letters and to the reviewer comments. If you are concerned about the release of confidential data, please indicate what specific information you would like to have removed; we cannot incorporate redactions for any other reasons. More information on transparent peer review is available.

Please do not hesitate to contact me should you have any questions.

Best wishes,

Pep

__
Pep Pàmies
Chief Editor, *Nature Biomedical Engineering*

P.S. Nature Portfolio journals encourage authors to share their step-by-step experimental protocols on a protocol-sharing platform of their choice. Nature Portfolio's Protocol Exchange is a free-to-use and open resource for protocols; protocols deposited in Protocol Exchange are citable and can be linked from the published article. More details can be found at www.nature.com/protocolexchange/about.

_____
Reviewer #1 (Report for the authors (Required)):

The authors have addressed my concerns.

Reviewer #2 (Report for the authors (Required)):

The title and some wording in the manuscript remains misleading and inaccurate in at least two ways:
1) The nucleases, base editors, and prime editors characterized in this study all use CRISPR components. So it is confusing and inaccurate to use "CRISPR" to mean "CRISPR nuclease". While there are base editors that do not use CRISPR, none of those variants are characterized in this study. So the authors should

replace “CRISPR” with “nuclease” or “Cas9 nuclease” or “CRISPR nuclease” in the title and in the manuscript.

2) The title remains misleading because it implies similar deletion outcomes when in fact they are 20-fold different! I do not know what the authors seem to remain fixed on this rather sensationalist and misleading title but the title of a scientific research paper should accurately summarize the key finding. This title does not. In fact, it implies the opposite of the primary finding in some respects. The title should be changed to accurately reflect that the nucleases in the study generated deletions at a 20-fold higher rather than the base or prime editors, consistent with previous reports.

_____

Rebuttal 1

Dear anonymous reviewers.

Thank you very much for the reviews and your comments about our manuscript (nBME-24-0092-T). Please find our detailed responses to the reviews shown in blue below. Revised sections of the main text are colored red in the revised manuscript. Thank you very much.

Sincerely yours,
Sangsu Bae, Ph. D.

Reviewers' comments:

Reviewer #1 (Report for the authors (Required)):

This manuscript by Gue-ho Hwang et al., entitled 'Detailed Mechanisms for Unintended Large DNA Deletions with CRISPR, Base Editors, and Prime Editors,' utilizes an optimized long-range amplicon sequencing system and develops a k-mer sequence-alignment algorithm to simultaneously detect small DNA alteration events and large DNA deletions. The authors determined the occurrence of large deletions in different cell lines and verified that TMEJ is the major process producing large deletions. Furthermore, using a similar method, they discovered large deletions in systems utilizing base editors and prime editors and proposed hypotheses to explain how small indels or large deletions occur in base editing systems.

Verifying the function of TMEJ represents an important discovery in this report; however, some observations presented in this manuscript have been similarly reported elsewhere. For example, the phenomenon of large deletions induced by CRISPR/Cas9 was already proposed by Fatwa Adikusuma et al. (PMID: 30089922), and similar hypotheses regarding indels in the base editing (BE) system have been mentioned elsewhere (PMID: 32690971; PMID: 27096365). Additionally, the occurrence of indels in H840A-nCas9, which is used in PE2 or PE3, has been previously proposed (PMID: 36997524). In addition, the nCas9 could induce indels mutation with two sgRNAs is also already confirmed (PMID: 23992846), thus two sgRNA might increasing the likelihood of large indels (>100bp, as mentioned in the manuscript) during the PE3 editing process.

Without a deeper clarification of the mechanism, retelling similar cases does not add novelty. In conclusion, I believe this article now does not meet the publication standards of NBE, and I suggest that more detailed mechanistic experiments should be conducted to enhance the value of the article.

Response: We really appreciate this reviewer's comprehensive summary with insightful critiques. We believe that our manuscript has been significantly improved during the revision process.

1. The term 'detailed mechanisms' has been proposed in the article title to describe the workings of 'CRISPR, base editors, and prime editors.' However, upon review, I found that the detailed mechanism of large deletions was only discussed in the context of CRISPR. For the base editor section, the mechanism of insertions and deletions (indels) was presented

merely as a hypothesis by the author. Moreover, the 'detailed mechanism' behind the occurrence of large indels in prime editing has not been clearly elucidated in this manuscript. Moreover, POLQ significantly contributes to the repair of double-strand breaks via the theta-mediated end joining (TMEJ) pathway, however, might not be key regulator in the gene editing in base editors or primer editors, for these system relies on nickase Cas9. Therefore, the term 'detailed mechanism' might only be applicable to CRISPR, but not to base editing or prime editing, as presented in this manuscript. If the authors still wish to use this term, further experiments are needed to more thoroughly clarify the mechanisms by which large indels occur in both base editing and prime editing.

Response: As the reviewer pointed out, we conducted CRISPRi screening experiments only for CRISPR-Cas9 nucleases and mainly elucidated the detailed mechanism for the CRISPR-Cas9 nucleases. Therefore, following the reviewer's suggestion, we decided to generally tone down the term "detailed mechanism" overall and changed the title to "Unintended large DNA deletions occur through cellular DNA repair pathways during gene editing with CRISPR, base editors, and prime editors".

On the other hand, we have enhanced the mechanism study with BE and PE during this revision process, which includes i) measuring large deletion events in various KO cell lines (*LIG IV* KO, *POLQ* KO, *RAD52* KO HeLa cell line) with BE and PE, ii) the treatment of Pol θ inhibitor (ART558) in addition to ATR inhibitor (M4344) for Cas9, BE, PE, and iii) double nick experiments for understanding the major factor of PE3 platform. We will discuss them in below.

2. In the result section "Optimizing a long-range amplicon sequencing method to determine small indels and large deletions with a high accuracy", The author highlights the high accuracy of the Illumina platform, which justifies the initial use of large fragment PCR followed by the development of specific algorithms for indel matching base on NGS platform. However, when it comes to the methodology for constructing long fragments, the approach lacks significant innovation. Conducting PCR on a 10kb genomic length is relatively straightforward. For example, in the full-length mitochondrial genome off-target detection process (PMID: 32641830), after long range PCR amplication (2 PCR, total nearly 16kb), by employing ultrasonication for random fragmentation followed by library construction and sequencing, it is feasible to assemble the sequences and identify indels or SNVs across the entire fragment. This technique is already well-established in the field. Consequently, I find that this method does not introduce notable innovation in the context of existing techniques.

From an algorithmic perspective, the author should highlight the advancements in accuracy and convenience offered by the newly established algorithm, rather than primarily emphasizing its "high speed and less memory" in the main text and Figure 1a. In the context of detecting indels within a 10kb DNA sequence, if indexes are exclusively created for 10Kb segments, utilizing BWA for alignment and subsequent analysis allows for rapid completion, without significantly taxing hardware memory. On the other hand, the claims of "high speed and less memory" made by the author require validation through comparative analysis with other algorithms.

Additionally, the reliability of this algorithm is fundamental to the entire article. Therefore,

using limited endogenous sites (only two, in this manuscript) to validate the reliability of a method is insufficient. The author needs more evidence to demonstrate that this newly established algorithm can effectively search for large deletions.

In addition, the reliable of the k-mer_align tool should be compare with other software or algorithm. For instance, previous algorithms have been shown in literature not to produce significant large indels in the BE system (PMID: 32711379), yet the results in this article suggest the opposite. The algorithm may be a crucial factor contributing to these differing outcomes. If the author need to be more invincible, I believe further parallel validation of the algorithm is essential. If this point are not proved well, then in every subsequent result, I recommend adding steps to validate with other algorithms (possibly in the supplementary material) to confirm the reliability of the findings.

Furthermore, the author could synthetically create or establish more complex library sequences, including both small indels and large deletions, to elucidate this issue in depth, making your results more credible.

Response: Thank you for valuable comments here. For several comments raised by the reviewer, we would like to address one by one as follows.

1) We also recognize that the long-range amplicon sequencing is a pre-exiting technique to analyze long DNA sequences. As the reviewer mentioned, full-length mitochondrial genome was sequenced by the Liu group in the previous paper (PMID: 32641830). However, please note that the aim of our study is different from that paper. The length-dependence analysis or measurement of exact ratios of different genome sizes were not considered in the previous paper. For example, if there are various mitochondrial genome sequences of different sizes, the shorter genome may be enriched during the PCR process. In our experiments, we optimized the long-range amplicon sequencing method and developed a dedicated program, named “Expanded Cas-Analyzer (ExCas-Analyzer)”.

2) There are three important improvements to optimize the long-range amplicon sequencing method: i) selecting an optimized DNA polymerase to minimize length bias during PCR, ii) developing a specialized k-mer alignment algorithm for our purpose, and iii) implementing strategies to eliminate background noise from false-positive large deletions, as shown in revised supplementary Figure 3.

3) As the reviewer suggested, we compared ExCas-Analyzer program with bwa mem that is a widely used alignment program and with CRISPResso2 that is a widely used analysis program for CRISPR. To this end, we produced the simulation data set for the long-range amplicon sequencing. ExCas-Analyzer has higher analysis speed and uses lower usage of memory than bwa and CRISPResso2. Additionally, we confirmed the reliability of ExCas-Analyzer using a simulated long-range sequencing data set. We generated simulation data containing large deletions at rates of 10%, 20%, 30% and 40% using the ART NGS simulator [PMID: 22199392] and checked how close ExCas-Analyzer came to the expected ratio. The table below shows the results, and R^2 was 0.8506, verifying the reliability of ExCas-Analyzer. Now, we are developing a web tool for the NGS data for Long-range amplicon sequencing to allow more researchers to freely use this method. And the fast analysis and low usage of memory will be a powerful advantage in developing web tool. We have added these results in revised Supplementary Fig. 6.

Added Supplementary Figure 6e.

Expected ratio	Proposed ratio
10%	11.82%
20%	23.13%
30%	37.14%
40%	43.26%

Program	Run time (sec)	Maximum usage of memory (MB)
ExCas-Analyzer	2.93	17
bwa mem	10.63	90
CRISPResso2	86.91	1499

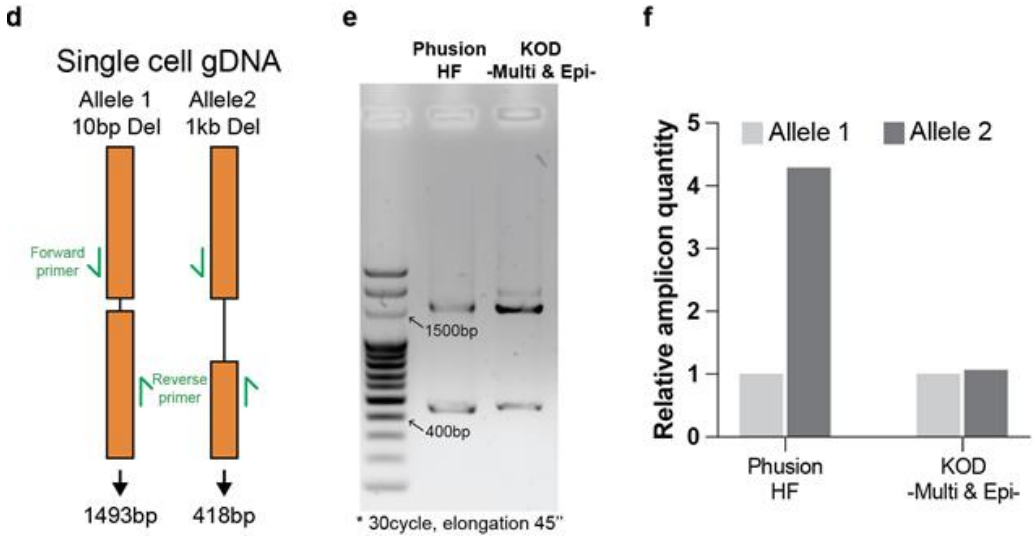
3. The preference of PCR enzymes is important, but I noticed that the gradient of preference used by the author is: 120, 140, 160, 180, 200, 220, 240, 260, 280bp. Given that the author's goal involves large fragments of 15kb and the focus is on large deletions (> 100bp), this gradient does not significantly contribute evidence towards enzyme preference for this article. It is necessary to include fragments with greater differences, such as 100bp, 500bp, 1000bp, and 10000bp, to conduct relevant experiments and further substantiate the author's point. Moreover, the extension time should be tested when performing PCR on longer sequences since it is possible that the extension time could be a key factor in limiting the length of the PCR product and may contribute to bias in the length of the PCR product.

Additionally, for the author's experiment, the ability to successfully perform PCR on 10kb genomic fragments is also very important. Evidence (gel running) of successful long range PCR with these five enzymes needs to be provided (which can be included in the supplementary material).

Response: Thank you for the comments. As the reviewer pointed out, in initial experiments, we used mixtures of short fragments because it is convenient to prepare the fragments having same primer binding sequences in equimolar amounts. And we also thought that the PCR bias would be conserved between mixtures in short range from 120 bp to 280 bp and mixtures in long range from 100bp to 10000bp.

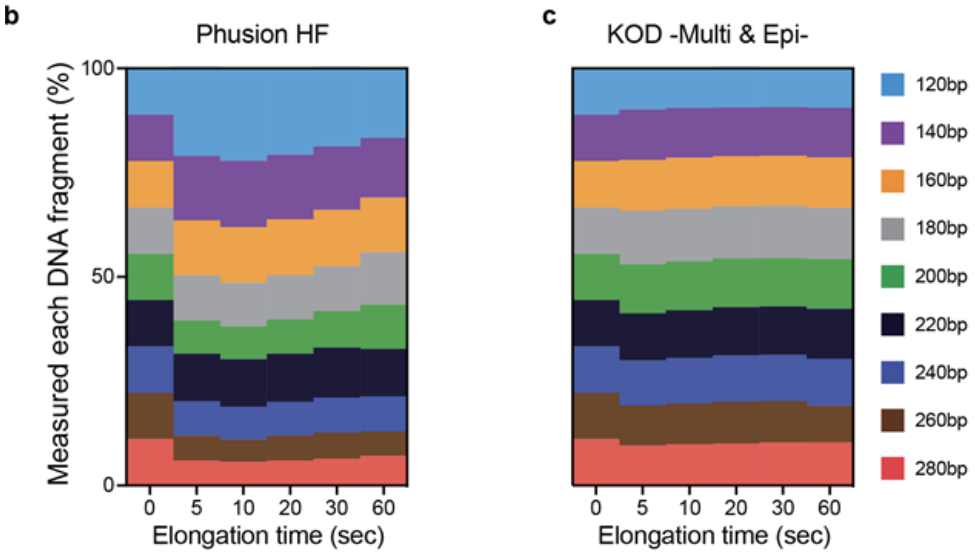
To address the reviewer's concern, we used a single cell line in which genomic DNA has 10bp deletion on one allele and 1kb deletion on the other allele. Through PCR amplification with a same primer set, 1493bp size of PCR product is amplified from 10bp-deletion-allele and 418bp size of PCR product is amplified from 1kb-deletion-allele. We tested two representative DNA polymerases – i) Phusion HF and ii) KOD -Multi & Epi- DNA polymerase which were determined to have high and low length bias, respectively. After experiment of agarose gel electrophoresis, we found that Phusion HF has more severe length bias compared to the KOD -Multi & Epi- DNA polymerase, consistent with the previous results with mixture from 120 bp to 280 bp. We have added these results to Supplementary Figure 4. From these data, we confirmed that the characteristics of PCR length bias are conserved to fragments in both short range and long range.

Added Supplementary Figure 4d-4f.



On the other hand, we further invested whether the PCR elongation time can affect the PCR length bias with the mixture of small fragments. Results showed that PCR elongation time is crucial for the PCR length bias, but Phusion HF created much larger PCR length bias than KOD - Multi & Epi- DNA polymerase even in saturated time (>40 sec). Please note that elongation time was 45 sec, indicating that it was a enough time and thus, we can exclude the elongation time effect. We have added these results to Supplementary Figure 4. For 10-15kb PCR amplification, elongation time (>6 min) was sufficiently long, according to the manufacturer's protocol.

Added Supplementary Figure 4b-4c.



4. In lines 167-169, it is observed that the CRISPRi library used by the authors specifically targeted 794 genes referenced in Repair-seq. However, in Figure 3B, only the genes associated with End protection and End resection are presented. This raises a question: what about the other genes? It is difficult to believe that only the Z-scores of End protection and End resection have undergone changes. What if there are additional genes that play a

crucial role in DSB processes? To address this concern, it would be valuable for the authors to include a figure displaying the Z-scores of other significantly altered genes not belonging to the End protection and End resection categories. Furthermore, it is important for the authors to engage in a discussion regarding these findings to fully explore the potential implications.

Response: Thank you for valuable comments. As the reviewer suggested, we further analyzed to find other genes that may affect the generation of large deletions. By comparing the p-value of CRISPR-treated data with the data of negative control in each gene, we identified top ten genes associated with an increase/decrease in large deletions and added the results to revised Supplementary Figure 9.

Added Supplementary Fig. 9.

a

Gene	P-value	Target_1	Target_2	Target_3	Gene Annotation
H2AFX	0.005409	-1.72	-1.64	-2.03	Histone H2A
RNF8	0.005409	-1.64	-1.63	-1.7	E3 ubiquitin-protein ligase RNF8
SNW1	0.010954	-1.99	-1.95	-1.07	SNW domain-containing protein 1
GADD45A	0.011497	-1.19	-1.53	-1.85	Growth arrest and DNA damage-inducible protein GADD45 alpha
GTF3C4	0.013275	-1.84	-1.63	-0.93	General transcription factor 3C polypeptide 4
HDAC2	0.013275	-2.14	-0.87	-1.94	Histone deacetylase 2
MRE11A	0.015289	-2.28	-0.75	-2.54	Double-strand break repair protein MRE11
NBN	0.017571	-2.49	-1.23	-1.03	Nibrin
COPS6	0.017941	-2.48	-2.47		COP9 signalosome complex subunit 6
RHNO1	0.018393	-1.87	-1.25	-0.96	RAD9, HUS1, RAD1-interacting nuclear orphan protein 1

b

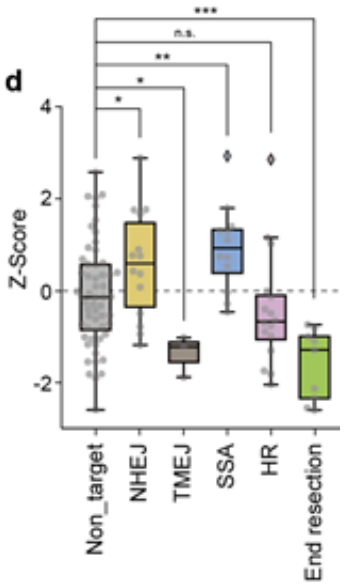
Gene	P-value	Target_1	Target_2	Target_3	Gene Annotation
CDYL	0.011497	2.08	1.93	1.36	Chromodomain Y-like protein
CHD2	0.014591	1.58	1.86	1.79	Down syndrome cell adhesion molecule
CREBBP	0.019252	1.06	2.82	1.24	CREB-binding protein
SIRT6	0.020075	2.01	2.7		NAD-dependent protein deacylase sirtuin-6
UHRF1	0.02304	1.11	1	2.08	E3 ubiquitin-protein ligase UHRF1
POLE3	0.025165	1.37	0.67	2.41	DNA polymerase epsilon subunit 3
UBE2B	0.025165	0.99	1.78	1.45	Ubiquitin-conjugating enzyme E2 B
NHEJ1	0.027459	1.16	0.7	4.39	Non-homologous end-joining factor 1
RFWD3	0.02993	2.47	2.38	0.35	E3 ubiquitin-protein ligase RFWD3
UBE2T	0.02993	1.01	0.87	2.04	Ubiquitin-conjugating enzyme E2 T

Interestingly, from the CRISPRi screening experiment, we found several genes associated with the increase or decrease of large deletion frequencies, in addition to the DNA repair pathway-associated genes. For example, UBE2T coordinates repair pathway in Fanconi anaemia (FA) pathway that is one of DNA damage response60, UBE2B is involved in monoubiquitylating histone when replication forks encounter lesions in post-replication repair pathway, UHRF1 plays a crucial role in the DNA damage response during the S phase, RFWD3 interacts with RPA to facilitate the DNA damage response, and RNF8 ubiquitylates histones as part of the DNA damage response. In this study, however, we focused on the DNA repair pathway-associated genes but further studies such as CRISPRi screening in genome level would be necessary to fully understand the factors contributing to the generation of large deletions in Cas9, BEs, and PEs. We have added this description to the discussion section [lines 385-399].

5. The selection of candidate genes depicted in Figure 3e for investigating the major pathway linked to large deletions raises some questions. For instance, it is notable that ERCC4, a gene within the SSA gene group, was not included as a candidate gene. Furthermore, there is a lack of gene knockout experiments conducted within the HR gene group. It is crucial to provide a clear explanation for the reasons behind these selection and experimental decisions. If you claim that another pathway is not the major pathway for large deletions to occur in gene editing, it is necessary to provide direct evidence to support this exclusion.

Response: As the reviewer commented, it would be best to measure the changes of large deletion frequencies in many genes for each repair pathway, as possible. However, please understand that major experiment of this study is the CRISPRi screening, and we believe that KO cell lines can be used for verification. Moreover, as shown in Figure 3d, we could not see any significant change (i.e., p-value was n.s.) in HR, thus we excluded to construct HR-related gene KO cell lines. We believe that CRISPRi screening as itself can be a direct evidence for that TMEJ is a major pathway.

Main Fig. 3d.

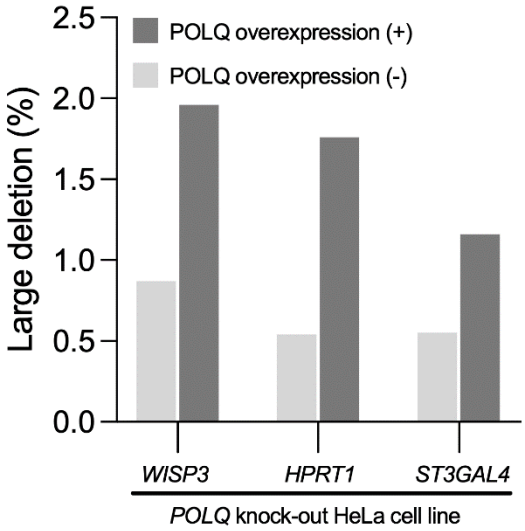


6. Figure 3e demonstrates that gene silencing of PQLQ could significantly decrease the occurrence of large deletions, implying that TMEJ is the primary pathway involved in large deletions. To further support this conclusion, it is essential to confirm whether overexpression of PQLQ could increase the incidence of large deletions in CRISPR gene editing. This experiment would add substantial strength to the author's argument and help to further establish the role of TMEJ in the mechanism of large deletions.

Response: Thank you for the comments. As the reviewer suggested, we examined whether the POLQ overexpression could increase the large deletion frequencies in *POLQ* knock-out cell line. Results exhibited that the large deletion frequencies were increased for targets in three independent genes, confirming that the TMEJ is the primary pathway for large deletion.

We have added the results to revised Supplementary Figure 8.

Added Supplementary Fig. 8b. POLQ overexpression experiments in *POLQ* knock-out cell line.



7. In lines 185-188, the authors stated, "Using a previously defined classification scheme, we categorized the genes into those associated with NHEJ (LIG4, XRCC4, XRCC5, and XRCC6), end resection (MRE11A, NBN, and RAD50), TMEJ (POLQ), SSA (ERCC4 and RAD52), and HR (BRCA1, BRCA2, RAD51AP1, RAD51B, RAD51C, and RAD51D)." From this perspective, individually knocking out each of these genes might be a more suitable approach than CRISPRi screening.

Response: We think this comment is similar to above comment #5.

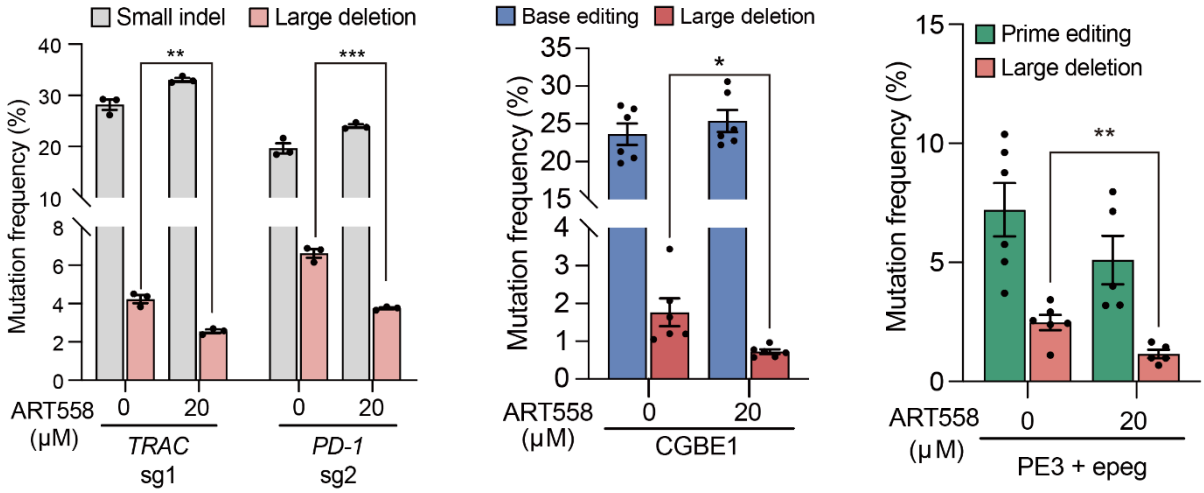
As the reviewer commented, it would be best to measure the changes of large deletion frequencies in many genes for each repair pathway, as possible. However, please understand that major experiment of this study is the CRISPRi screening, and we believe that KO cell lines can be used for verification. Moreover, as shown in Figure 3d, we could not see any significant change (i.e., p-value was n.s.) in HR, thus we excluded to construct HR-related gene KO cell lines. We believe that CRISPRi screening as itself can be a direct evidence for that TMEJ is a major pathway.

8. In line 232, the author mentioned “Collectively, these data support evaluation of M4344 or drugs that reduce end resection or TMEJ pathways as co-effectors with CRISPR-Cas9-mediated genetic engineering in therapeutic applications to mitigate unintended large deletions”. It’s a good point that that M4344 or drugs targeting the end resection or TMEJ pathways could potentially be used alongside CRISPR-Cas9-mediated genetic engineering to mitigate unintended large deletions in therapeutic applications. However, it should be noted that the primary function of M4344 is to pharmacologically inhibit ATR activity, thereby suppressing the entire end resection process, including TMEJ, SSA, and HR. In this study, the author did not employ any specific strategy to individually inhibit the TMEJ pathway, such as POLQ inhibitor (ART558) in primary T cells.

Response: Thank you for valuable comments here. To address the reviewer's concern, in addition to M4344, we applied POLQ inhibitor (ART558) to investigate whether the generation of undesired large deletion was mitigated in HeLa cell line. We found that, similar to M4344, ART558 reduced undesired Cas9-mediated large deletion as well (added to revised Fig. 3g). As the reviewer mentioned, because the M4344 can affect the entire end resection process, the ART558 may be more suitable as a large deletion inhibitor.

In addition, we also applied ART558 to base editor (CGBE1) and prime editor (PE3). Results exhibited that ART558 effectively mitigated the undesired large deletion in both CGBE1 (added to revised Fig. 4f) and PE3 (added to revised Fig. 5f). Taken together, we suggest that ART558 can be a suitable candidate to mitigate undesired large deletions in therapeutic applications with Cas9, base editor and prime editor. We have added these results in revised Figure 3g, Figure 4f and Figure 5f, respectively.

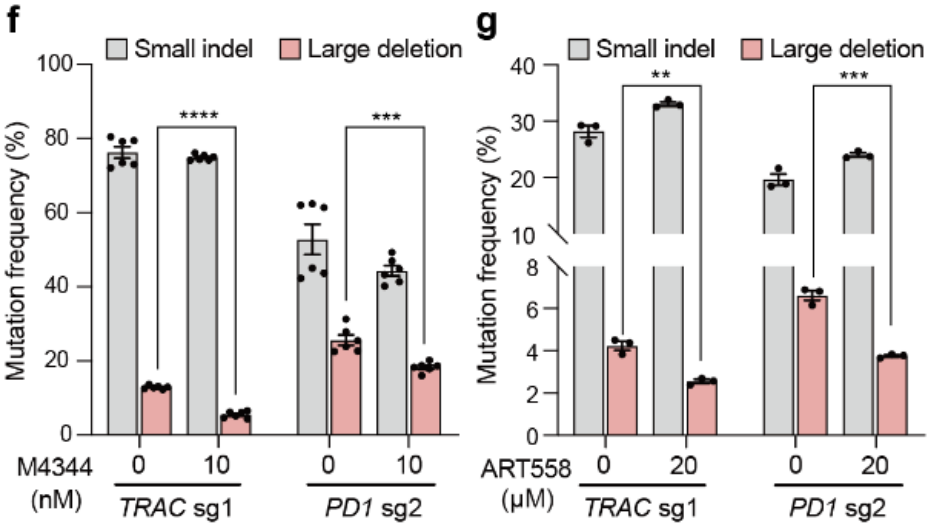
Added Figure 3g, 4f, and 5f.



9. In the section titled "Confirmation by knockout that end resection and TMEJ processes generate large deletions", I observed that the author solely focuses on the decreased occurrence of large deletions during the gene editing process. However, it is crucial to also examine the on-target efficiency of the M4344+ or M4344- groups in the tested samples. Was there any impact of M4344 on the normal editing efficiency of the target genes? If the on-target effect was indeed reduced, it would consequently lead to a decrease in off-target effects as well. The relationship between on-target efficiency and the occurrence of large deletions is a crucial factor to consider when determining the suitability of an assistant drug in gene therapy. To provide a comprehensive understanding, it is important for the author to include information about the on-target effects in the same figure. This would allow the readers to discern that the decrease in large deletions is not solely attributed to a decline in the editing efficiency of the gene editor.

Response: Thank you for valuable comments here. To address the reviewer's concern, we investigated the small indel and large deletion mutations with/without M4344 and also with/without ART558. We found that there were no significant decrease of small indel frequencies regardless of the M4344 or ART558 treatments (p-value = 0.3972 for *TRAC* sg1, 0.0787 for *PD-1* sg2). We have added these results in revised Figure 3f and 3g.

Revised Figure 3f-3g



10. In line 526, the author mentioned "HeLa CRISPRi cells were generated by lentiviral integration (~3 to 5 MOI) using the dCas9-KRAB-blast plasmid (Addgene #89567), followed by single cell isolation." However, this raises a question about the rationale behind the single cell isolation. It seems the procedure involves introducing a lentivirus library into the cells before isolating individual cells. If this is indeed the case, it implies the potential creation of over 795 unique cell lines, each corresponding to different single-guide RNAs (sgRNAs) for CRISPR interference (CRISPRi). Such a scenario suggests that 795 individual transfections of ribonucleoprotein complexes (RNPs) would be necessary for the screening experiment. Nevertheless, the text does not mention any details regarding the execution of these individual transfections, leaving a gap in the explanation of the methodology.

Furthermore, the methodology for screening sgRNAs and confirming their association with the signaling pathway needs to be meticulously detailed in the methods section. It is equally important that a comprehensive description of these processes is presented in the results section to ensure clarity and understanding.

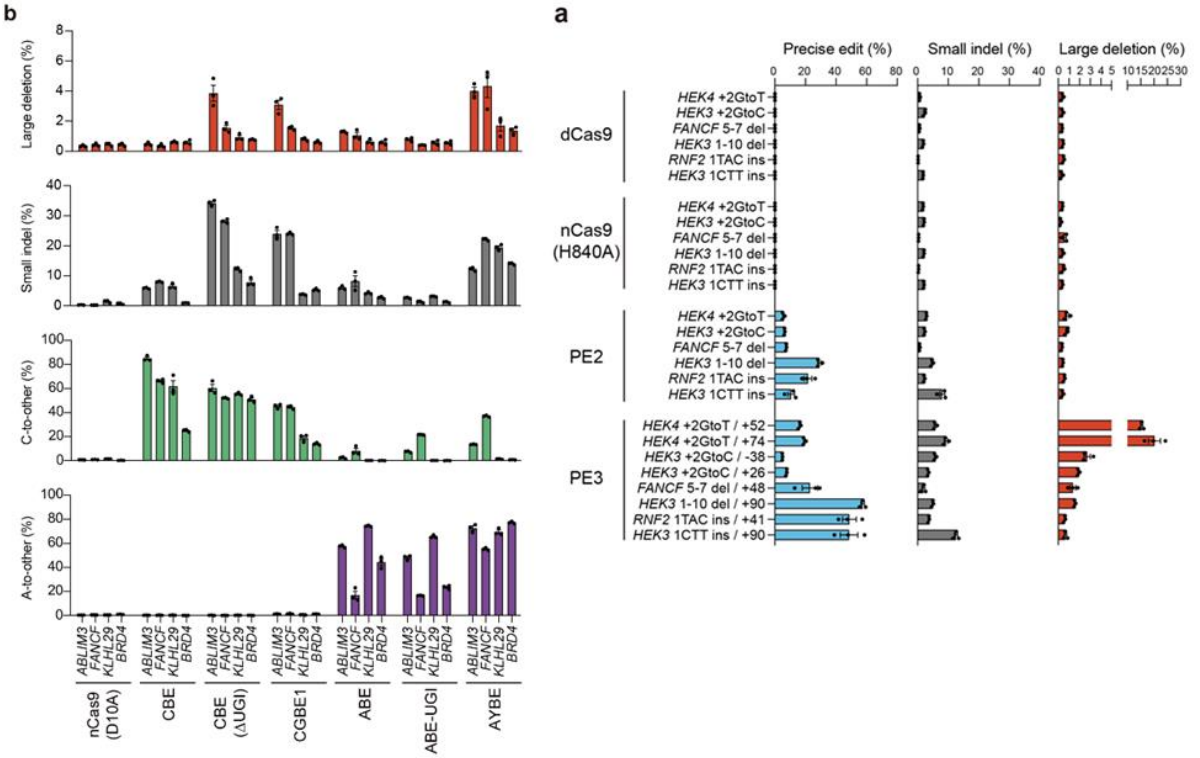
Response: We apology for the ambiguous description, potentially leading to misunderstandings. In fact, single cell was isolated to establish CRISPRi HeLa single cell line, after lenti-viral transduction with the CRISPRi library. To clarify this, we added a dedicated section to clarify the exact protocol for CRISPRi HeLa cell line construction in [lines 594-598].

11. It is important to note that both base editors and prime editors utilize nCas9 as the underlying structure. In this manuscript, the base editors mentioned employ D10A nCas9 as the base architecture, while PE2 and PE3 use H840A nCas9. Considering this perspective, it is recommended to include a control group with nCas9 in Figure 4 and Figure 5, targeting different sites.

Response: Thank you for valuable comments here. To address the reviewer's concern, we conducted all experiments with nCas9(D10A) as a BE control and with nCas9(H840A) as a

PE control. The updated data with nCas9(D10A) and nCas9(H840A) were added to revised Figure 4b-4d and 5a-5b, respectively.

Revised Figure 4b (left) and 5a (right)



12. In Figure 4, it is essential to provide a detailed version and plasmid information of the cytosine base editors (CBE), cytosine-guanine base editors (CGBE), adenine base editors (ABE), and adenine-to-cytosine base editors (AYBE).

Response: We have added all specific version of BEs in main text (lines 269-286) and plasmid information in the method section (lines 658-662).

13. In Figure 4e, the author presents a hypothesis concerning the occurrence of small indels or large deletions. However, it is important to note that similar hypothesis or part of the hypothesis has already been proposed in previous reports (PMID: 32690971; PMID: 27096365). Thus, the hypothesis lacks novelty.

Response: As the reviewer commented, the two papers suggest similar model with us. However, as the reviewer mentioned in above Q2, some studies did not see significant large deletions in the BE system [PMID: 32711379], whereas some observed BE-mediated small indels or large deletions [PMID: 32690971; PMID: 27096365], which is controversial. In our study, we concluded that BEs are reasonable for large deletion generation, especially for BE platforms that easily produce AP sites [i.e., CBE(ΔUGI), CGBE1, AYBE]. Based on our observation, we clarify BE-mediated DSBs and large deletion with a Schematic.

14. As mentioned earlier, there have been reports suggesting that base editing does not result in large deletions. It is commendable that the author has identified instances of large deletions occurring in base editors. However, the conflicting viewpoints need to be carefully addressed and discussed. To strengthen the argument, more compelling evidence and better discussion should be provided.

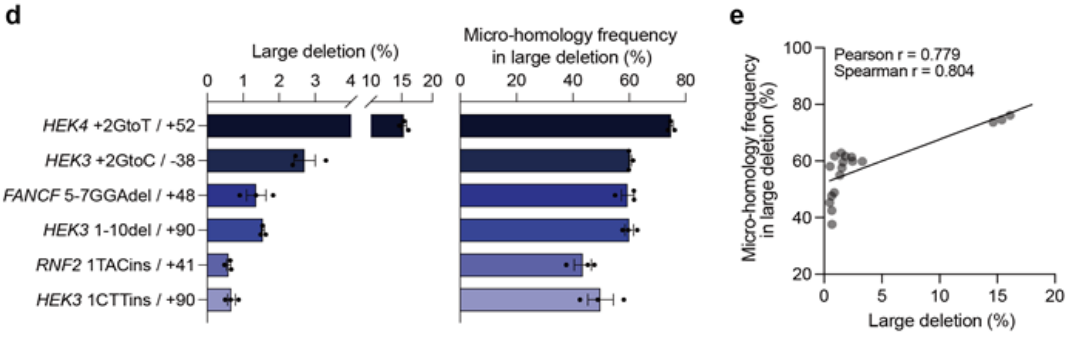
Response: To the best of our knowledge, no one have tested with various BE platforms to investigate the occurrence of DNA large deletion events. In the present study, we used 6 representing BE forms and found that BE platforms that easily produce AP sites [i.e., CBE(ΔUGI), CGBE1, AYBE] have tendencies to generate large deletion more, for the first time. KO cell line test in Figure 4e and ART treatment experiment in Figure 4f further enhance our findings.

15. In Figure 5C, we observe that the efficiency of large deletions using PE3 is categorized into two distinct groups: one group exhibits an efficiency of approximately 3%, while the other demonstrates a significantly higher efficiency around 50%. This substantial variance warrants further discussion to understand the underlying causes.

Moreover, it is important to explore the potential relationship between these higher editing efficiency and the large deletion efforts of PE2/PE3.

Response: Thank you for the comments. In fact, it is not easy to determine why PE3 generates different outcomes of small indels and large deletions in different target sites. Regarding the reviewer's comment, we speculated that the sequence context including microhomology information might be one reason. Thus, we investigated the microhomology (2-8bp) frequencies within large deletion patterns and compared the large deletion frequencies with microhomology frequencies. We observed a positive correlation between them. Therefore, we assumed that the target dependency could be one reason for generating large deletion events. We have added these results to revised Supplementary Fig. 12.

Revised Supplementary Figure 12d-12e.

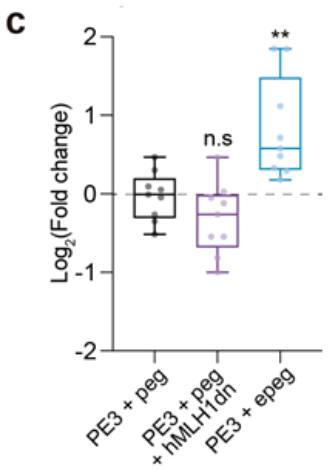


16. The results of PE3+epeg, as shown in Figure 5d, exhibit slightly significant variability. To obtain more reliable data, increasing the sample size is recommended.

Response: As the reviewer suggested, we conducted the experiments with additional epegRNAs (total 3 independent targets) and replicates (n=3). To examine the effect of

hMLH1dn and epegRNA, we displayed the data with fold change values of large deletion frequency comparing to the frequency of PE3+peg approach. We have updated data in revised Figure 5c.

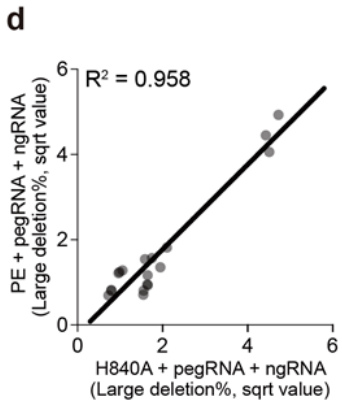
Revised Figure 5c



17. In Fig 5 and discussion section, further discussion is warranted regarding the underlying mechanism by which the PE3 system leads to a higher incidence of indels and large deletions. Previous studies have demonstrated that efficient targeting effects can be achieved through the combined action of two nCas9 enzymes (PMID: 23992846). As the PE3 system employs a dual sgRNA strategy, which shares similarities with the principle of gene knockout using nickase Cas9, it is crucial to include a control group in Figure 5. Specifically, adding nCas9 + sgRNA1 + sgRNA2 (where sgRNA1/2 corresponds to the sgRNAs utilized in PE3) might help ascertain whether the observed large indels are solely attributable to nCas9.

Response: Thank you for valuable comments here. As the reviewer suggested, we conducted nCas9 (H840A) with pegRNA and ngRNA. When we treated nCas9 (H840A) with pegRNA and ngRNA, instead of PE2 with pegRNA and ngRNA (i.e., without RT) in human cells, we found a strong correlation of large deletion frequencies between them, indicating that the generation of double nicks by nCas9 is the most critical factor for generating DSBs and large deletions in PE3. We have added this data in revised Figure 5d.

Revised Figure 5d.

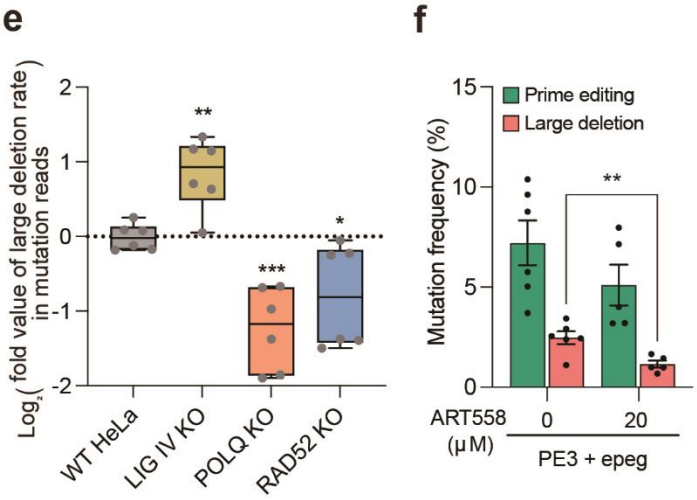


18. In Figure 5, it is evident that the authors have not provided mechanistic explanations or a hypothesis. This lack of information creates a disconnect from the "detailed mechanism" (in title) concerning prime editors. Therefore, it is crucial to design and incorporate additional experiments pertaining to the mechanism in Figure 5.

Response: As the reviewer mentioned, we generally toned down the term "detailed mechanism" overall and changed the title accordingly. However, we did our best to reveal the mechanism of PE-mediated indels and large deletion, as follows.

We further investigated subsequent pathways leading to large deletions with PE3 platform in various KO cell lines (*LIG IV* KO, *POLQ* KO, *RAD52* KO HeLa cell line) used in Figure 3e. Similar to Cas9 nucleases and BEs, the frequency of large deletions was increased in *LIG IV* KO cell line, but decreased in *POLQ* KO cell line (**Fig. 5e**). Also, the treatment of ART558 decreased the large deletion events (**Fig. 5f**), suggesting that TMEJ remains the major pathway for PE-mediated large deletions. However, we also observed that the frequency of large deletions was decreased in *RAD52* KO cell line, opposite to Cas9 nucleases (**Fig. 5e**), possibly implying the difference between DNA repair processes after Cas9- and PE-driven DSBs.

Revised Figure 5e-5f.



Reviewer #2 (Report for the authors (Required)):

The manuscript "Detailed mechanisms for unintended large DNA deletions with CRISPR, base editors, and prime editors," by Hwang, Lee, and colleagues sought to characterize the frequencies of and mechanisms by which large DNA deletion byproducts are generated during the course of genome editing experiments. To address the quantification problem, they developed a long-range next-generation sequencing (NGS) pipeline to investigate the genotypes that arise from Cas9 nuclease cutting, base editing, and prime editing. To investigate the mechanism by which these genotypes occur, they performed a CRISPRi screen and found that long deletion genotypes arise from the theta-mediated end-joining (TMEJ) pathway. Lastly, they examined a variety of base editing architectures and prime editing systems to investigate the frequencies of long deletions in those systems.

The work addresses an important question in the application of therapeutic genome editing: "How can one quantify and reduce undesired genotypes that cannot be sequenced directly with traditional NGS techniques?" However, while this question is important, there are several aspects of this work that must be addressed before it is suitable for publication in Nature Biomedical Engineering. Addressing the following points would significantly strengthen the manuscript, and improve its relevance to a biomedical engineering audience:

Response: We appreciate this reviewer's supportive and positive comments.

Major points:

- (Lines 38–39): The maximum frequencies of the deletions that the authors found for biomedically relevant base editors (CBE: 0.8%, ABE: 1.42%) are quite small. Moreover, the frequencies of these deletions are low for prime editors as well, with a maximum large-deletion frequency of 1.2% for PE2 and a clear bimodal distribution for PE3, with the majority of edits having a large-deletion frequency under 5%. The abstract as written makes a misleading and false equivalence between Cas9-based and nCas9-based technologies as generating similar levels of such byproducts. In reality, the main finding, that BE and PE result in far lower frequencies of these large deletions, is less sensationalist but much more important and accurate. The authors should reword this sentence to make this distinction more clear or remove this sentence entirely from the abstract, as it is misleading and sensationalist.

Response: We appreciate the reviewer's comment. As suggested, we have revised the abstract with the expression "fewer but significant large deletions compared to Cas nucleases" to avoid the misunderstanding in [lines 40-41].

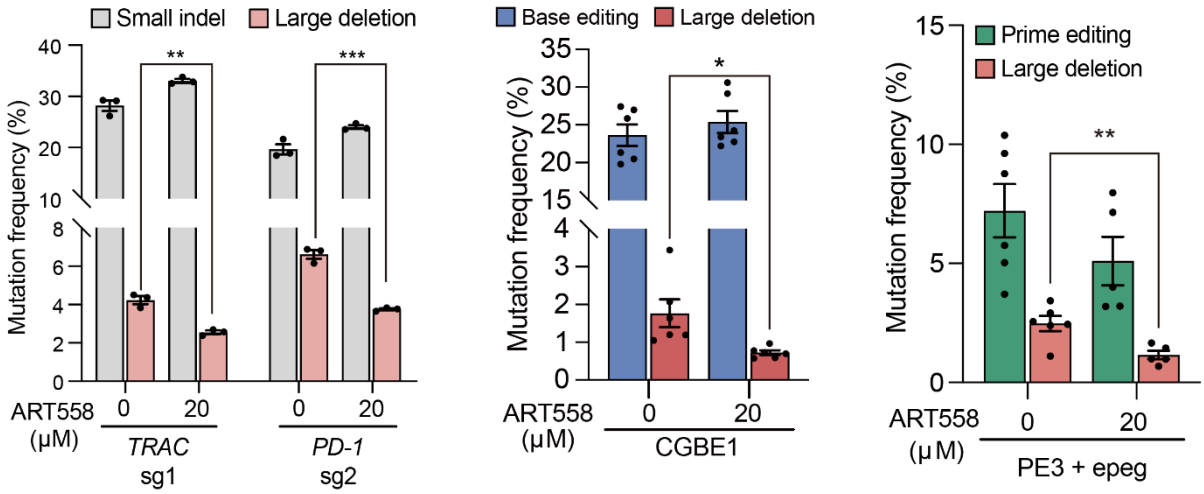
- (232–234): While M4344 works for reducing TMEJ-mediated large deletions, inhibition of ATR kinase would also preclude the HDR mechanism and would therefore be unsuitable for approaches used in many applications of Cas9 nuclease, such as integration of a chimeric antigen receptor into T cells. The authors should make this caveat more clear. Alternatively, a more convincing argument could be made using an inhibitor of DNA polymerase theta to selectively inhibit TMEJ but not HDR.

Response: Thank you for valuable comments here. To address the reviewer's concern, in

addition to M4344, we applied POLQ inhibitor (ART558) to investigate whether the generation of undesired large deletion was mitigated in HeLa cell line. We found that, similar to M4344, ART558 reduced undesired Cas9-mediated large deletion as well (added to revised Fig. 3g). As the reviewer mentioned, because the M4344 can affect the entire end resection process, the ART558 may be more suitable as a large deletion inhibitor.

In addition, we also applied ART558 to base editor (CGBE1) and prime editor (PE3). Results exhibited that ART558 effectively mitigated the undesired large deletion in both CGBE1 (added to revised Fig. 4f) and PE3 (added to revised Fig. 5f). Taken together, we suggest that ART558 can be a suitable candidate to mitigate undesired large deletions in therapeutic applications with Cas9, base editor and prime editor. We have added these results in revised Figure 3g, Figure 4f and Figure 5f, respectively.

Added Figure 3g, 4f, and 5f.



- (249–253): Importantly, CBE(ΔUGI), CGBE, and AYBE are not relevant to clinical applications of base editing—they are explicitly avoided for clinical use because their mechanism yields byproducts including deletions (though still fewer than nucleases). A better way of structuring this section is to present the canonical CBE/ABE results (which show low deletion rates), and then describe the additional non-clinical constructs tested as part of a mechanistic analysis showing how these low-frequency deletions arise. Then, the authors can add that based on their analysis, ABE–UGI could be used in cases where there are TC motifs present in the base editing window to further drive down indels resulting from bystander cytidine deamination.

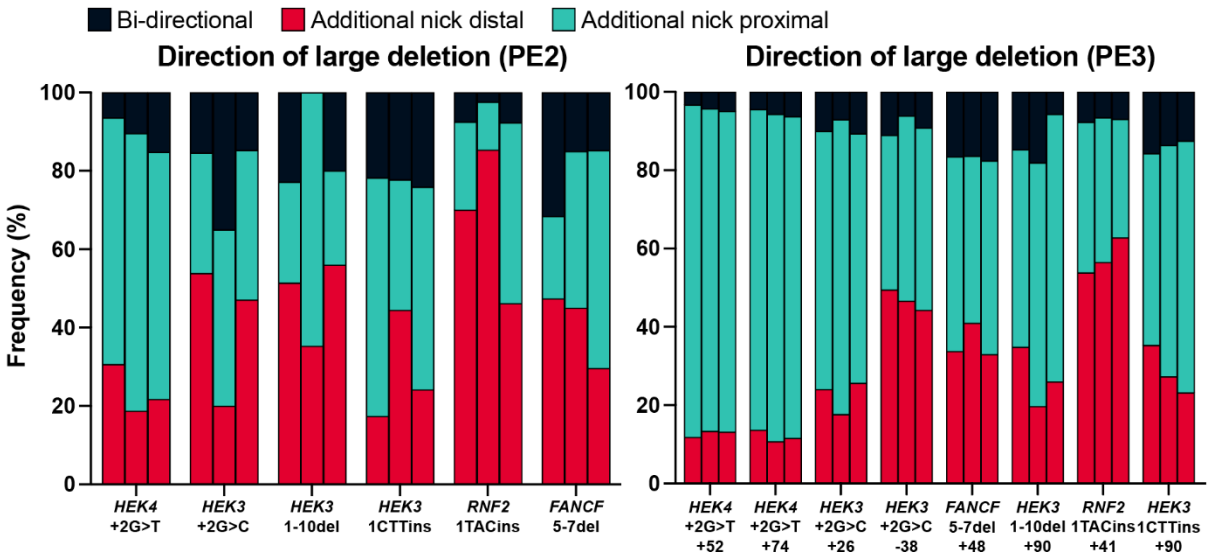
Response: Thank you for the comments. We agree with the comments that the canonical CBE or ABE have much chance to be used for further clinical applications because of low byproducts compared to CBE(ΔUGI), CGBE, and AYBE. However, in terms of expanding the knowledge of BE platforms, we believe that various BE platforms including CBE(ΔUGI), CGBE, and AYBE should be compared in parallel. Moreover, someone may improve the CGBE or AYBE as a safety form in the future. Also, we found an effort to utilize GBE (Glycosylase base editor) as a gene editing therapeutic application named *BRL-103* (ClinicalTrials.gov identifier; NCT05442346). Please understand the current structure.

- (293–294): The authors should examine the exact deletion genotypes that arise from PE3. Do these deletions extend past the two nick sites? If so, this would suggest a mechanism that also involves resection mediated by ATR kinase and therefore treatment with M4344 should reduce the frequency of these deletion genotypes. If not, then the authors should note that the mechanism by which these genotypes occur is distinct from that discussed earlier in the paper, and perhaps propose a different mechanism that can explain the results observed.

Response: Thank you for the comments.

To address the reviewer's concern, we analyzed the direction of large deletion events generated by PE2 or PE3. Basically, the origin of large deletion would be the nick site by pegRNA. The directions can be divided into three groups (Bi-directional, additional nick distal and additional nick proximal). We determined directionality when one side is twice longer than the other, and judged proximity based on the additional nick site. In case of PE3, the additional nick proximal large deletion patterns are increased compared to the PE2. We have added the data to revised Supplementary Figure 12.

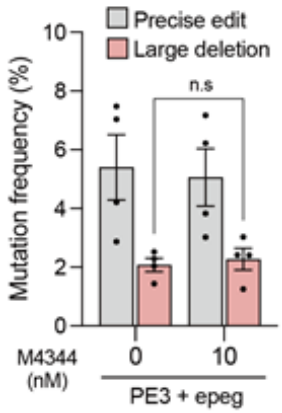
Added Supplementary Figure 12a.



We further examined whether M4344 can reduce the large deletion frequency or not. Results showed that M4344 did not contribute to reduce the large deletion frequency in PE3 approach, in contrast to the POLQ inhibitor (ART558) that reduce the large deletion frequency well. We have added the data to revised Supplementary Figure 12.

Added Supplementary Figure 12f.

f



When we treated PE3 platform in various KO cell lines (LIG IV KO, POLQ KO, RAD52 KO HeLa cell line), we found that the frequency of large deletions was increased in LIG IV KO cell line, but decreased both in POLQ KO cell line and in RAD52 KO cell line, which is different from the case of Cas9 nucleases. Taken together, we concluded that TMEJ remains the major pathway for PE-mediated large deletions but detailed mechanisms might be different between DNA repair processes after Cas9- and PE-driven DSBs.

- (452–496): Were read pairs deduplicated before processing? This step is crucial for quantitative analyses involving DNA shearing followed by a barcoding amplification step because PCR bias during the barcoding step can result in some sequences being duplicated and therefore overrepresented.

Response: Thank you for comments. In this study, we developed optimized long-range amplicon sequencing method and applied it to most experiments without the CRISPRi screening experiment. For the optimized long-range amplicon sequencing method, there are three important improvements to mitigate PCR-bias: i) selecting an optimized DNA polymerase to minimize length bias during PCR, ii) developing a specialized k-mer alignment algorithm for our purpose, and iii) implementing strategies to eliminate background noise from false-positive large deletions, as shown in revised supplementary Figure 3. Although we did not employ Unique Molecular Identifiers (UMIs), we confirmed the accuracy of our protocol and pipeline in different ways as shown in Figure 1.

Minor points:

- (136, 486): In the main text, the authors defined the cutoff for large vs. small deletions at 100 bp. However, in the methods, the authors defined the cutoff at 50 bp. Can the authors please clarify this discrepancy?

Response: Thank you for the comment. We fixed it to 100 bp.

- (318–322): New data should not be included in the discussion section. Instead, the authors can consider making a new results sub-section to discuss mischaracterization of cell lines. It would also strengthen the overall impact of the paper if the authors can make an analysis of the likelihood of such mischaracterization in current clinically relevant genome editing workflows.

Response: Thank you for the comment. For the previous Supplementary Figure 10, we have moved it to revised Supplementary Figure 3. But for the previous Supplementary Figure 11, we decided to remove it in the revised manuscript because we think this data is somewhat redundant.

- (452–496): The authors did not specify the k-mer length they used for analyses performed in this work. Additionally, the use of the letter k for both the sequence length in the k-mer hash table and the k-means clustering hyperparameter is confusing.

Response: Thank you for comments. We have clarified the words. “k” for k-mer hash was to k_{hash} and “k” for k-means clustering was to $k_{cluster}$. We used k_{hash} for 11.

Reviewer #3 (Report for the authors (Required)):

This is a nice article in which the authors highlight the issue of long deletions (>100 bp) after gene editing. Although the long deletion phenomenon has been previously described (as authors cite), here the authors develop some improved experimental and analytic assays for long-range amplification, fragmented short-read sequencing or long-read sequencing, and edit sequence alignment, conduct a genetic screen identifying end resection as an important pathway required for long deletions, and demonstrate the risks of long deletions with certain base and prime editing approaches. This will be a useful reference for future investigators seeking to comprehensively characterize gene editing outcomes including long deletion events and is a cautionary tale that long deletions may indeed be present after various forms of gene editing if only one looks carefully.

Response: We appreciate this reviewer's supportive and positive comments.

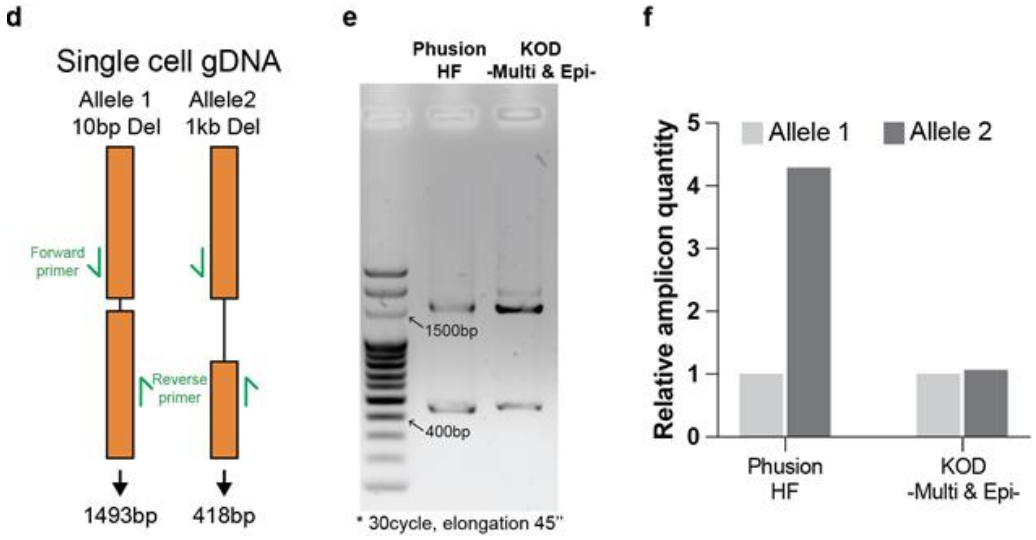
Some specific comments:

1. Fig 1c shows minimal bias comparing a ~9kb and ~10kb product (~10% difference in length). Also Supp Fig 3 is a nice design, but only tests range of 120-280 bp. What about longer deletions, e.g. up to 9 kb deletions for a 10 kb product (~90% difference in length)? It would be interesting to see the performance of the polymerases around a range of larger deletions to better understand possible length biases of the procedures.

Response: Thank you for the comments. As the reviewer pointed out, in initial experiments, we used mixtures of short fragments because it is convenient to prepare the fragments having same primer binding sequences in equimolar amounts. And we also thought that the PCR bias would be conserved between mixtures in short range from 120 bp to 280 bp (2.3-fold difference here) and mixtures in long range from 1kb to 10kb

To address the reviewer's concern, we used a single cell line in which genomic DNA has 10bp deletion on one allele and 1kb deletion on the other allele. Through PCR amplification with a same primer set, 1493bp size of PCR product is amplified from 10bp-deletion-allele and 418bp size of PCR product is amplified from 1kb-deletion-allele (3.6-fold difference here). When we tested two representative DNA polymerases – i) Phusion HF and ii) KOD -Multi & Epi- DNA polymerase which were determined to have high and low length bias, respectively, we found that the KOD -Multi & Epi- DNA polymerase still showed less bias, consistent with the results with mixture from 120 bp to 280 bp. We have added these results to Supplementary Figure 4.

Added Supplementary Figure 4d-4f.



2. Regarding the performance of the alignment strategies, it could be interesting to simulate reads with various deletions and compare approaches.

Response: Thank you for valuable comments here. As the reviewer suggested, we compared ExCas-Analyzer program with bwa mem that is a widely used alignment program and with CRISPResso2 that is a widely used analysis program for CRISPR. To this end, we produced the simulation data set for the long-range amplicon sequencing. ExCas-Analyzer has higher analysis speed and uses lower usage of memory than bwa and CRISPResso2. Additionally, we confirmed the reliability of ExCas-Analyzer using a simulated long-range sequencing data set. We generated simulation data containing large deletions at rates of 10%, 20%, 30% and 40% using the ART NGS simulator [PMID: 22199392] and checked how close ExCas-Analyzer came to the expected ratio. The table below shows the results, and R^2 was 0.8506, verifying the reliability of ExCas-Analyzer. Now, we are developing a web tool for the NGS data for Long-range amplicon sequencing to allow more researchers to freely use this method. And the fast analysis and low usage of memory will be a powerful advantage in developing web tool. We have added these results in revised Supplementary Fig. 6.

Added Supplementary Figure 6e.

Expected ratio	Proposed ratio
10%	11.82%
20%	23.13%
30%	37.14%
40%	43.26%

Program	Run time (sec)	Maximum usage of memory (MB)
ExCas-Analyzer	2.93	17
bwa mem	10.63	90
CRISPResso2	86.91	1499

3. In Fig 2, are the small insertions templated +1 position 17 reiterations or short homonucleotide insertions as have been described as a common Cas9 repair outcome (PMC5834694) or some other type of small insertion which might reflect a different repair mechanism? The anti-correlation might suggest competing DNA repair mechanisms.

Response: As the reviewer suggested, we investigated the insertion patterns for *HPRT1, ST3GAL4,* and *WISP3* in HeLa and HEK293T cell lines. Consistent with the findings reported in the study [PMC5834694], the predominant insertion pattern observed was a 1nt homonucleotide insertion, as illustrated in the figure below. For insertions longer than 1bp, the sequence adjacent to the insertion site was often repeated.

Figure for the reviewer #3 only. Various outcomes of insertion pattern after Cas9 treated.

HeLa

HPRT1

CTTGAGCACAC	C	AGAGGGCTA	68.87
TTGAGCACAC	A	AGAGGGCTAC	6.47
CTTGAGCACAC	T	AGAGGGCTA	4.91
CCCTTGAGCACAC	AC	AGAGG	4.7
CCCTTGAGCACAC	ACAC	AGAGG	1.02

ST3GAL4

CCACGACCACA	A	CAGCGGCGG	91.15
TCCCCACGACCACA	CA	CAGCGG	1.14
CCACGGCCACA	A	CAGCGGCGG	0.39
CCACGACCACA	G	CAGCGGCGG	0.39
TCCCCACGACCACA	AA	CAGCGG	0.38

WISP3_sg1

CCAGGAGGGCA	A	ACGGGGCTT	88.9
CCAGGAGGGCA	C	ACGGGGCTT	0.85
CCAGGAGGGCA	G	ACGGGGCTT	0.64
CCAGGAGGGCA	A	ACGGGGCTC	0.53
CCAGGAGGGCA	A	GCGGGGCTT	0.46

HEK293T

HPRT1

CTTGAGCACAC	C	AGAGGGCTA	56.97
TTGAGCACAC	A	AGAGGGCTAC	4.85
CTTGAGCACAC	T	AGAGGGCTA	3.64
CTTGAGCACAC	AC	AGAGG	3.64
CTTGAGCACAC	G	AGAGGGCTA	2.42

ST3GAL4

CCACGACCAC	A	ACAGCGGCGG	87.21
CCACGACCAC	AC	ACAGCGG	1.21
CCACGACCAC	AG	CAGCGGCGG	0.76
CACGACCAC	CTG	AGCGGCGGC	0.45
CCACGACCAC	C	ACAGCGGCG	0.38

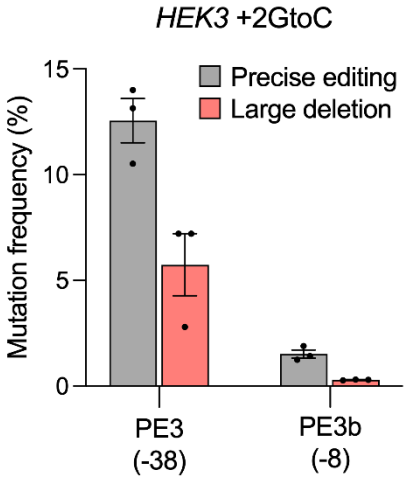
WISP3_sg1

CCAGGAGGGCA	A	ACGGGGCTT	89.88
CAGGAGGGCA	A	ACGGGGCTTC	0.58
CCAGGAGGGCA	AA	ACGGGGCTT	0.52
CCAGGAGGGCA	T	ACGGGGCTT	0.52
CCAGGAGGGCA	G	ACGGGGCTT	0.23

4. Fig 5 and Supp Fig 9: it could be interesting to try PE3b strategy to see if it is more similar to PE2 or PE3 in terms of long deletions.

Response: Thank you for the comments. As the reviewer suggested, we compared PE3 and PE3b platforms with *HEK3* +2GtoC epegRNA. Results showed that large deletion frequency of PE3b was quite decreased compared to the value of PE3. However, the prime editing frequency was also decreased because of different additional nicking gRNA. Theoretically, the PE3b may reduce the frequency of large deletions since the double nicks did not occur simultaneously, consistent with the decreased occurrence of small indels with PE3b. But, we have very few cases, we have just added description about it in the discussion section [lines 385-387].

Figure for the reviewer #3 only. Comparison between PE3 and PE3b.



5. Could the authors add some discussion regarding to what degree including a UMI might increase the accuracy and minimize the bias of the long read quantification?

Response: Thank you for the comments. In this study, we developed optimized long-range amplicon sequencing method and applied it to most experiments without the CRISPRi screening experiment. For the optimized long-range amplicon sequencing method, there are three important improvements to mitigate PCR-bias: i) selecting an optimized DNA polymerase to minimize length bias during PCR, ii) developing a specialized k-mer alignment algorithm for our purpose, and iii) implementing strategies to eliminate background noise from false-positive large deletions, as shown in revised supplementary Figure 11. Although we did not employ Unique Molecular Identifiers (UMIs), we confirmed the accuracy of our protocol and pipeline in different ways as shown in Figure 1.

However, as the reviewer suggested, long-read sequencing platform with a UMI may aid to detect large deletion. We have added description about it in the discussion section [lines 368-370].

6. A few typos:

- Supp Fig1b "bluw" typo.

- Typo line 261-2 "In Figure 4c, each large deletion frequency was divided by the average large deletion frequency with."

- Fig S10 typo "Stard"

- For lines 321 and 322, is it meant "homozygous" and "heterozygous" rather than "homologous" and "heterologous"?

Thank you for the comments. We have fixed all in the revised manuscript.

Rebuttal 2

Dear anonymous reviewers.

Thank you very much for the reviews and your comments about our manuscript (nBME-24-0092-T). Please find our detailed responses to the reviews shown in blue below. Revised sections of the main text are colored red in the revised manuscript. Thank you very much.

Sincerely yours,
Sangsu Bae, Ph. D.

Reviewers' comments:

Reviewer #1 (Report for the authors (Required)):

This manuscript by Gue-ho Hwang et al., entitled 'Detailed Mechanisms for Unintended Large DNA Deletions with CRISPR, Base Editors, and Prime Editors,' utilizes an optimized long-range amplicon sequencing system and develops a k-mer sequence-alignment algorithm to simultaneously detect small DNA alteration events and large DNA deletions. The authors determined the occurrence of large deletions in different cell lines and verified that TMEJ is the major process producing large deletions. Furthermore, using a similar method, they discovered large deletions in systems utilizing base editors and prime editors and proposed hypotheses to explain how small indels or large deletions occur in base editing systems.

Verifying the function of TMEJ represents an important discovery in this report; however, some observations presented in this manuscript have been similarly reported elsewhere. For example, the phenomenon of large deletions induced by CRISPR/Cas9 was already proposed by Fatwa Adikusuma et al. (PMID: 30089922), and similar hypotheses regarding indels in the base editing (BE) system have been mentioned elsewhere (PMID: 32690971; PMID: 27096365). Additionally, the occurrence of indels in H840A-nCas9, which is used in PE2 or PE3, has been previously proposed (PMID: 36997524). In addition, the nCas9 could induce indels mutation with two sgRNAs is also already confirmed (PMID: 23992846), thus two sgRNA might increasing the likelihood of large indels (>100bp, as mentioned in the manuscript) during the PE3 editing process.

Without a deeper clarification of the mechanism, retelling similar cases does not add novelty. In conclusion, I believe this article now does not meet the publication standards of NBE, and I suggest that more detailed mechanistic experiments should be conducted to enhance the value of the article.

Response: We really appreciate this reviewer's comprehensive summary with insightful critiques. We believe that our manuscript has been significantly improved during the revision process.

1. The term 'detailed mechanisms' has been proposed in the article title to describe the workings of 'CRISPR, base editors, and prime editors.' However, upon review, I found that the detailed mechanism of large deletions was only discussed in the context of CRISPR. For the base editor section, the mechanism of insertions and deletions (indels) was presented

merely as a hypothesis by the author. Moreover, the 'detailed mechanism' behind the occurrence of large indels in prime editing has not been clearly elucidated in this manuscript. Moreover, POLQ significantly contributes to the repair of double-strand breaks via the theta-mediated end joining (TMEJ) pathway, however, might not be key regulator in the gene editing in base editors or primer editors, for these system relies on nickase Cas9. Therefore, the term 'detailed mechanism' might only be applicable to CRISPR, but not to base editing or prime editing, as presented in this manuscript. If the authors still wish to use this term, further experiments are needed to more thoroughly clarify the mechanisms by which large indels occur in both base editing and prime editing.

Response: As the reviewer pointed out, we conducted CRISPRi screening experiments only for CRISPR-Cas9 nucleases and mainly elucidated the detailed mechanism for the CRISPR-Cas9 nucleases. Therefore, following the reviewer's suggestion, we decided to generally tone down the term "detailed mechanism" overall and changed the title to "Unintended large DNA deletions occur through cellular DNA repair pathways during gene editing with CRISPR, base editors, and prime editors".

On the other hand, we have enhanced the mechanism study with BE and PE during this revision process, which includes i) measuring large deletion events in various KO cell lines (*LIG IV* KO, *POLQ* KO, *RAD52* KO HeLa cell line) with BE and PE, ii) the treatment of Pol θ inhibitor (ART558) in addition to ATR inhibitor (M4344) for Cas9, BE, PE, and iii) double nick experiments for understanding the major factor of PE3 platform. We will discuss them in below.

2. In the result section "Optimizing a long-range amplicon sequencing method to determine small indels and large deletions with a high accuracy", The author highlights the high accuracy of the Illumina platform, which justifies the initial use of large fragment PCR followed by the development of specific algorithms for indel matching base on NGS platform. However, when it comes to the methodology for constructing long fragments, the approach lacks significant innovation. Conducting PCR on a 10kb genomic length is relatively straightforward. For example, in the full-length mitochondrial genome off-target detection process (PMID: 32641830), after long range PCR amplication (2 PCR, total nearly 16kb), by employing ultrasonication for random fragmentation followed by library construction and sequencing, it is feasible to assemble the sequences and identify indels or SNVs across the entire fragment. This technique is already well-established in the field. Consequently, I find that this method does not introduce notable innovation in the context of existing techniques.

From an algorithmic perspective, the author should highlight the advancements in accuracy and convenience offered by the newly established algorithm, rather than primarily emphasizing its "high speed and less memory" in the main text and Figure 1a. In the context of detecting indels within a 10kb DNA sequence, if indexes are exclusively created for 10Kb segments, utilizing BWA for alignment and subsequent analysis allows for rapid completion, without significantly taxing hardware memory. On the other hand, the claims of "high speed and less memory" made by the author require validation through comparative analysis with other algorithms.

Additionally, the reliability of this algorithm is fundamental to the entire article. Therefore,

using limited endogenous sites (only two, in this manuscript) to validate the reliability of a method is insufficient. The author needs more evidence to demonstrate that this newly established algorithm can effectively search for large deletions.

In addition, the reliable of the k-mer_align tool should be compare with other software or algorithm. For instance, previous algorithms have been shown in literature not to produce significant large indels in the BE system (PMID: 32711379), yet the results in this article suggest the opposite. The algorithm may be a crucial factor contributing to these differing outcomes. If the author need to be more invincible, I believe further parallel validation of the algorithm is essential. If this point are not proved well, then in every subsequent result, I recommend adding steps to validate with other algorithms (possibly in the supplementary material) to confirm the reliability of the findings.

Furthermore, the author could synthetically create or establish more complex library sequences, including both small indels and large deletions, to elucidate this issue in depth, making your results more credible.

Response: Thank you for valuable comments here. For several comments raised by the reviewer, we would like to address one by one as follows.

1) We also recognize that the long-range amplicon sequencing is a pre-exiting technique to analyze long DNA sequences. As the reviewer mentioned, full-length mitochondrial genome was sequenced by the Liu group in the previous paper (PMID: 32641830). However, please note that the aim of our study is different from that paper. The length-dependence analysis or measurement of exact ratios of different genome sizes were not considered in the previous paper. For example, if there are various mitochondrial genome sequences of different sizes, the shorter genome may be enriched during the PCR process. In our experiments, we optimized the long-range amplicon sequencing method and developed a dedicated program, named “Expanded Cas-Analyzer (ExCas-Analyzer)”.

2) There are three important improvements to optimize the long-range amplicon sequencing method: i) selecting an optimized DNA polymerase to minimize length bias during PCR, ii) developing a specialized k-mer alignment algorithm for our purpose, and iii) implementing strategies to eliminate background noise from false-positive large deletions, as shown in revised supplementary Figure 3.

3) As the reviewer suggested, we compared ExCas-Analyzer program with bwa mem that is a widely used alignment program and with CRISPResso2 that is a widely used analysis program for CRISPR. To this end, we produced the simulation data set for the long-range amplicon sequencing. ExCas-Analyzer has higher analysis speed and uses lower usage of memory than bwa and CRISPResso2. Additionally, we confirmed the reliability of ExCas-Analyzer using a simulated long-range sequencing data set. We generated simulation data containing large deletions at rates of 10%, 20%, 30% and 40% using the ART NGS simulator [PMID: 22199392] and checked how close ExCas-Analyzer came to the expected ratio. The table below shows the results, and R^2 was 0.8506, verifying the reliability of ExCas-Analyzer. Now, we are developing a web tool for the NGS data for Long-range amplicon sequencing to allow more researchers to freely use this method. And the fast analysis and low usage of memory will be a powerful advantage in developing web tool. We have added these results in revised Supplementary Fig. 6.

Added Supplementary Figure 6e.

Expected ratio	Proposed ratio
10%	11.82%
20%	23.13%
30%	37.14%
40%	43.26%

Program	Run time (sec)	Maximum usage of memory (MB)
ExCas-Analyzer	2.93	17
bwa mem	10.63	90
CRISPResso2	86.91	1499

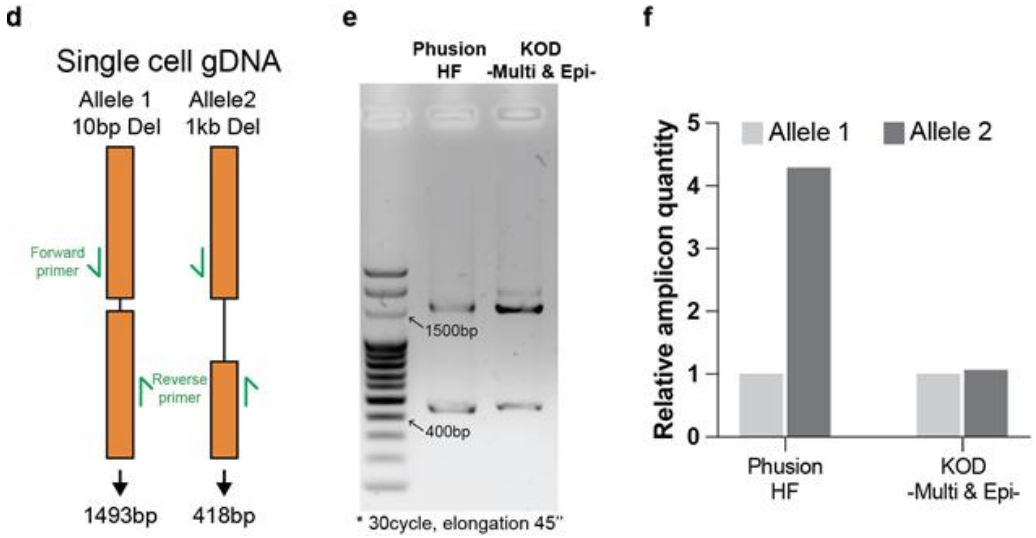
3. The preference of PCR enzymes is important, but I noticed that the gradient of preference used by the author is: 120, 140, 160, 180, 200, 220, 240, 260, 280bp. Given that the author's goal involves large fragments of 15kb and the focus is on large deletions (> 100bp), this gradient does not significantly contribute evidence towards enzyme preference for this article. It is necessary to include fragments with greater differences, such as 100bp, 500bp, 1000bp, and 10000bp, to conduct relevant experiments and further substantiate the author's point. Moreover, the extension time should be tested when performing PCR on longer sequences since it is possible that the extension time could be a key factor in limiting the length of the PCR product and may contribute to bias in the length of the PCR product.

Additionally, for the author's experiment, the ability to successfully perform PCR on 10kb genomic fragments is also very important. Evidence (gel running) of successful long range PCR with these five enzymes needs to be provided (which can be included in the supplementary material).

Response: Thank you for the comments. As the reviewer pointed out, in initial experiments, we used mixtures of short fragments because it is convenient to prepare the fragments having same primer binding sequences in equimolar amounts. And we also thought that the PCR bias would be conserved between mixtures in short range from 120 bp to 280 bp and mixtures in long range from 100bp to 10000bp.

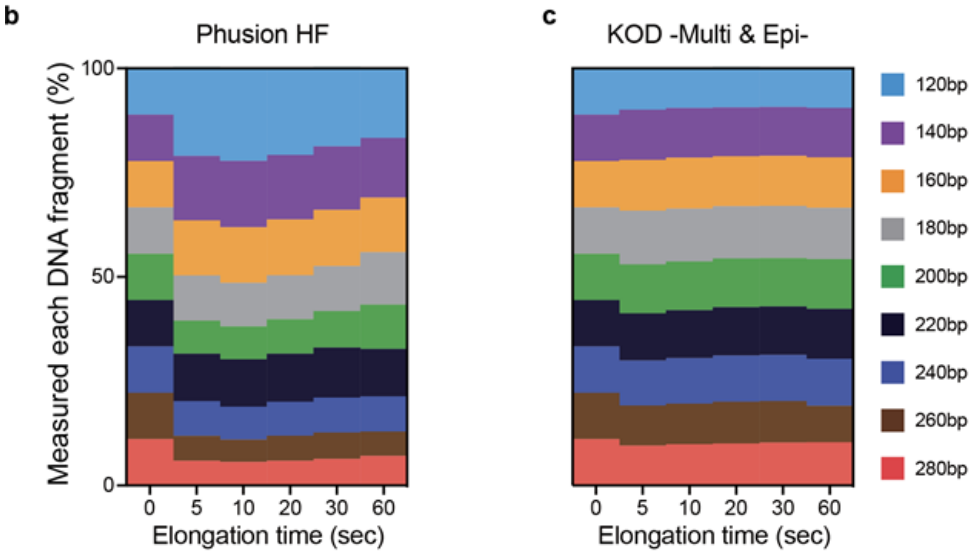
To address the reviewer's concern, we used a single cell line in which genomic DNA has 10bp deletion on one allele and 1kb deletion on the other allele. Through PCR amplification with a same primer set, 1493bp size of PCR product is amplified from 10bp-deletion-allele and 418bp size of PCR product is amplified from 1kb-deletion-allele. We tested two representative DNA polymerases – i) Phusion HF and ii) KOD -Multi & Epi- DNA polymerase which were determined to have high and low length bias, respectively. After experiment of agarose gel electrophoresis, we found that Phusion HF has more severe length bias compared to the KOD -Multi & Epi- DNA polymerase, consistent with the previous results with mixture from 120 bp to 280 bp. We have added these results to Supplementary Figure 4. From these data, we confirmed that the characteristics of PCR length bias are conserved to fragments in both short range and long range.

Added Supplementary Figure 4d-4f.



On the other hand, we further invested whether the PCR elongation time can affect the PCR length bias with the mixture of small fragments. Results showed that PCR elongation time is crucial for the PCR length bias, but Phusion HF created much larger PCR length bias than KOD - Multi & Epi- DNA polymerase even in saturated time (>40 sec). Please note that elongation time was 45 sec, indicating that it was a enough time and thus, we can exclude the elongation time effect. We have added these results to Supplementary Figure 4. For 10-15kb PCR amplification, elongation time (>6 min) was sufficiently long, according to the manufacturer's protocol.

Added Supplementary Figure 4b-4c.



4. In lines 167-169, it is observed that the CRISPRi library used by the authors specifically targeted 794 genes referenced in Repair-seq. However, in Figure 3B, only the genes associated with End protection and End resection are presented. This raises a question: what about the other genes? It is difficult to believe that only the Z-scores of End protection and End resection have undergone changes. What if there are additional genes that play a

crucial role in DSB processes? To address this concern, it would be valuable for the authors to include a figure displaying the Z-scores of other significantly altered genes not belonging to the End protection and End resection categories. Furthermore, it is important for the authors to engage in a discussion regarding these findings to fully explore the potential implications.

Response: Thank you for valuable comments. As the reviewer suggested, we further analyzed to find other genes that may affect the generation of large deletions. By comparing the p-value of CRISPR-treated data with the data of negative control in each gene, we identified top ten genes associated with an increase/decrease in large deletions and added the results to revised Supplementary Figure 9.

Added Supplementary Fig. 9.

a

Gene	P-value	Target_1	Target_2	Target_3	Gene Annotation
H2AFX	0.005409	-1.72	-1.64	-2.03	Histone H2A
RNF8	0.005409	-1.64	-1.63	-1.7	E3 ubiquitin-protein ligase RNF8
SNW1	0.010954	-1.99	-1.95	-1.07	SNW domain-containing protein 1
GADD45A	0.011497	-1.19	-1.53	-1.85	Growth arrest and DNA damage-inducible protein GADD45 alpha
GTF3C4	0.013275	-1.84	-1.63	-0.93	General transcription factor 3C polypeptide 4
HDAC2	0.013275	-2.14	-0.87	-1.94	Histone deacetylase 2
MRE11A	0.015289	-2.28	-0.75	-2.54	Double-strand break repair protein MRE11
NBN	0.017571	-2.49	-1.23	-1.03	Nibrin
COPS6	0.017941	-2.48	-2.47		COP9 signalosome complex subunit 6
RHNO1	0.018393	-1.87	-1.25	-0.96	RAD9, HUS1, RAD1-interacting nuclear orphan protein 1

b

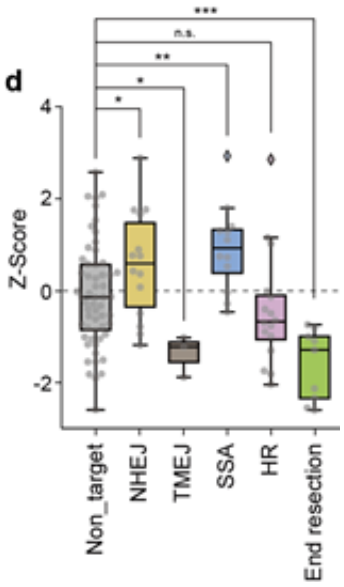
Gene	P-value	Target_1	Target_2	Target_3	Gene Annotation
CDYL	0.011497	2.08	1.93	1.36	Chromodomain Y-like protein
CHD2	0.014591	1.58	1.86	1.79	Down syndrome cell adhesion molecule
CREBBP	0.019252	1.06	2.82	1.24	CREB-binding protein
SIRT6	0.020075	2.01	2.7		NAD-dependent protein deacylase sirtuin-6
UHRF1	0.02304	1.11	1	2.08	E3 ubiquitin-protein ligase UHRF1
POLE3	0.025165	1.37	0.67	2.41	DNA polymerase epsilon subunit 3
UBE2B	0.025165	0.99	1.78	1.45	Ubiquitin-conjugating enzyme E2 B
NHEJ1	0.027459	1.16	0.7	4.39	Non-homologous end-joining factor 1
RFWD3	0.02993	2.47	2.38	0.35	E3 ubiquitin-protein ligase RFWD3
UBE2T	0.02993	1.01	0.87	2.04	Ubiquitin-conjugating enzyme E2 T

Interestingly, from the CRISPRi screening experiment, we found several genes associated with the increase or decrease of large deletion frequencies, in addition to the DNA repair pathway-associated genes. For example, UBE2T coordinates repair pathway in Fanconi anaemia (FA) pathway that is one of DNA damage response60, UBE2B is involved in monoubiquitylating histone when replication forks encounter lesions in post-replication repair pathway, UHRF1 plays a crucial role in the DNA damage response during the S phase, RFWD3 interacts with RPA to facilitate the DNA damage response, and RNF8 ubiquitylates histones as part of the DNA damage response. In this study, however, we focused on the DNA repair pathway-associated genes but further studies such as CRISPRi screening in genome level would be necessary to fully understand the factors contributing to the generation of large deletions in Cas9, BEs, and PEs. We have added this description to the discussion section [lines 385-399].

5. The selection of candidate genes depicted in Figure 3e for investigating the major pathway linked to large deletions raises some questions. For instance, it is notable that ERCC4, a gene within the SSA gene group, was not included as a candidate gene. Furthermore, there is a lack of gene knockout experiments conducted within the HR gene group. It is crucial to provide a clear explanation for the reasons behind these selection and experimental decisions. If you claim that another pathway is not the major pathway for large deletions to occur in gene editing, it is necessary to provide direct evidence to support this exclusion.

Response: As the reviewer commented, it would be best to measure the changes of large deletion frequencies in many genes for each repair pathway, as possible. However, please understand that major experiment of this study is the CRISPRi screening, and we believe that KO cell lines can be used for verification. Moreover, as shown in Figure 3d, we could not see any significant change (i.e., p-value was n.s.) in HR, thus we excluded to construct HR-related gene KO cell lines. We believe that CRISPRi screening as itself can be a direct evidence for that TMEJ is a major pathway.

Main Fig. 3d.

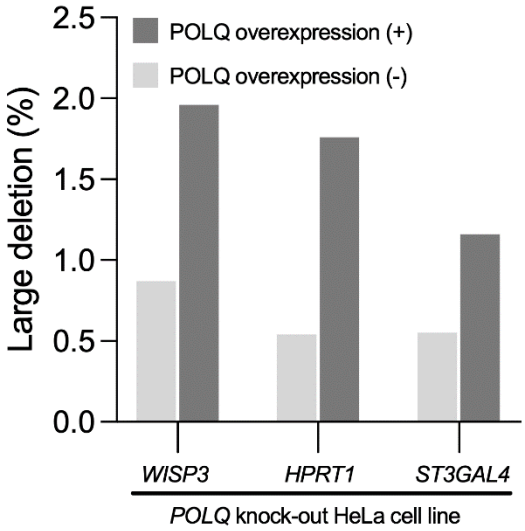


6. Figure 3e demonstrates that gene silencing of PQLQ could significantly decrease the occurrence of large deletions, implying that TMEJ is the primary pathway involved in large deletions. To further support this conclusion, it is essential to confirm whether overexpression of PQLQ could increase the incidence of large deletions in CRISPR gene editing. This experiment would add substantial strength to the author's argument and help to further establish the role of TMEJ in the mechanism of large deletions.

Response: Thank you for the comments. As the reviewer suggested, we examined whether the POLQ overexpression could increase the large deletion frequencies in *POLQ* knock-out cell line. Results exhibited that the large deletion frequencies were increased for targets in three independent genes, confirming that the TMEJ is the primary pathway for large deletion.

We have added the results to revised Supplementary Figure 8.

Added Supplementary Fig. 8b. POLQ overexpression experiments in *POLQ* knock-out cell line.



7. In lines 185-188, the authors stated, "Using a previously defined classification scheme, we categorized the genes into those associated with NHEJ (LIG4, XRCC4, XRCC5, and XRCC6), end resection (MRE11A, NBN, and RAD50), TMEJ (POLQ), SSA (ERCC4 and RAD52), and HR (BRCA1, BRCA2, RAD51AP1, RAD51B, RAD51C, and RAD51D)." From this perspective, individually knocking out each of these genes might be a more suitable approach than CRISPRi screening.

Response: We think this comment is similar to above comment #5.

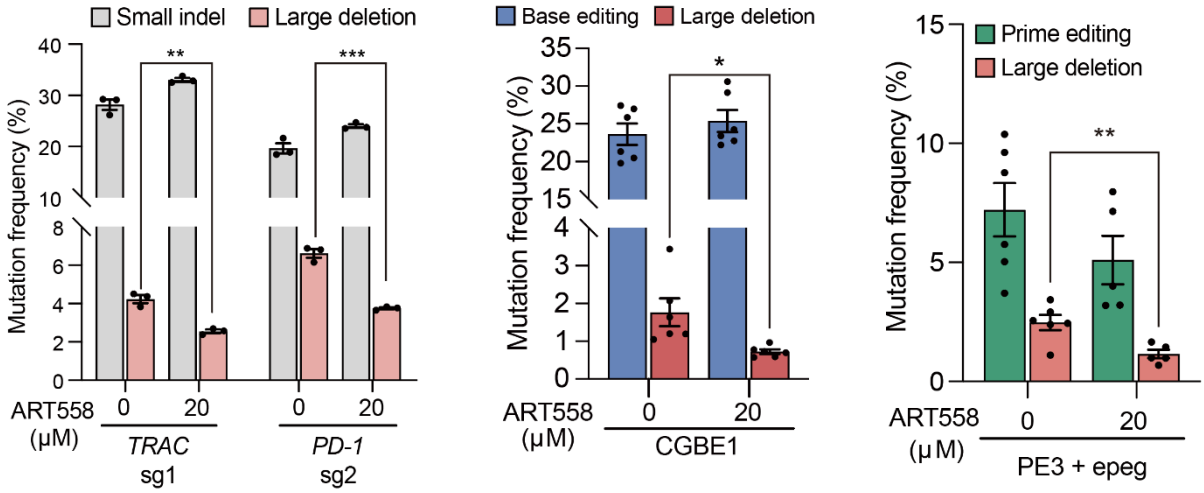
As the reviewer commented, it would be best to measure the changes of large deletion frequencies in many genes for each repair pathway, as possible. However, please understand that major experiment of this study is the CRISPRi screening, and we believe that KO cell lines can be used for verification. Moreover, as shown in Figure 3d, we could not see any significant change (i.e., p-value was n.s.) in HR, thus we excluded to construct HR-related gene KO cell lines. We believe that CRISPRi screening as itself can be a direct evidence for that TMEJ is a major pathway.

8. In line 232, the author mentioned “Collectively, these data support evaluation of M4344 or drugs that reduce end resection or TMEJ pathways as co-effectors with CRISPR-Cas9-mediated genetic engineering in therapeutic applications to mitigate unintended large deletions”. It’s a good point that that M4344 or drugs targeting the end resection or TMEJ pathways could potentially be used alongside CRISPR-Cas9-mediated genetic engineering to mitigate unintended large deletions in therapeutic applications. However, it should be noted that the primary function of M4344 is to pharmacologically inhibit ATR activity, thereby suppressing the entire end resection process, including TMEJ, SSA, and HR. In this study, the author did not employ any specific strategy to individually inhibit the TMEJ pathway, such as POLQ inhibitor (ART558) in primary T cells.

Response: Thank you for valuable comments here. To address the reviewer's concern, in addition to M4344, we applied POLQ inhibitor (ART558) to investigate whether the generation of undesired large deletion was mitigated in HeLa cell line. We found that, similar to M4344, ART558 reduced undesired Cas9-mediated large deletion as well (added to revised Fig. 3g). As the reviewer mentioned, because the M4344 can affect the entire end resection process, the ART558 may be more suitable as a large deletion inhibitor.

In addition, we also applied ART558 to base editor (CGBE1) and prime editor (PE3). Results exhibited that ART558 effectively mitigated the undesired large deletion in both CGBE1 (added to revised Fig. 4f) and PE3 (added to revised Fig. 5f). Taken together, we suggest that ART558 can be a suitable candidate to mitigate undesired large deletions in therapeutic applications with Cas9, base editor and prime editor. We have added these results in revised Figure 3g, Figure 4f and Figure 5f, respectively.

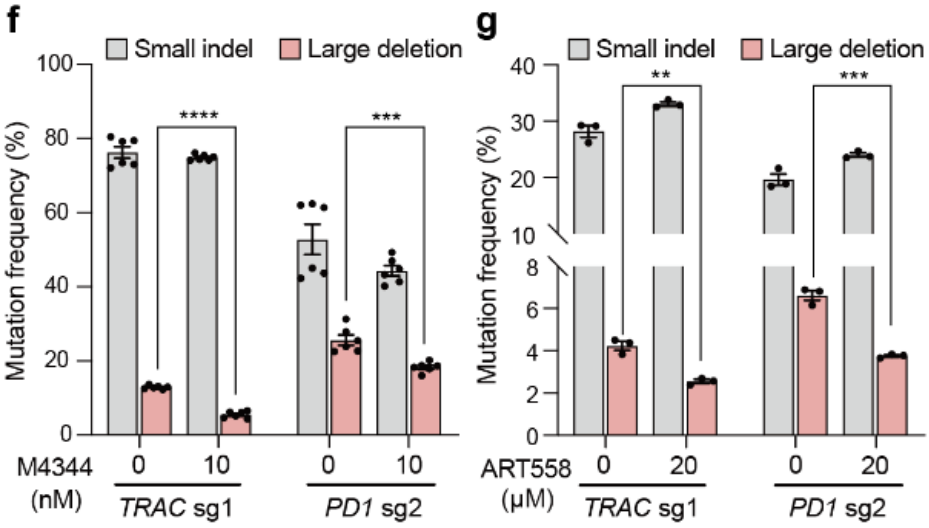
Added Figure 3g, 4f, and 5f.



9. In the section titled "Confirmation by knockout that end resection and TMEJ processes generate large deletions", I observed that the author solely focuses on the decreased occurrence of large deletions during the gene editing process. However, it is crucial to also examine the on-target efficiency of the M4344+ or M4344- groups in the tested samples. Was there any impact of M4344 on the normal editing efficiency of the target genes? If the on-target effect was indeed reduced, it would consequently lead to a decrease in off-target effects as well. The relationship between on-target efficiency and the occurrence of large deletions is a crucial factor to consider when determining the suitability of an assistant drug in gene therapy. To provide a comprehensive understanding, it is important for the author to include information about the on-target effects in the same figure. This would allow the readers to discern that the decrease in large deletions is not solely attributed to a decline in the editing efficiency of the gene editor.

Response: Thank you for valuable comments here. To address the reviewer's concern, we investigated the small indel and large deletion mutations with/without M4344 and also with/without ART558. We found that there were no significant decrease of small indel frequencies regardless of the M4344 or ART558 treatments (p-value = 0.3972 for *TRAC* sg1, 0.0787 for *PD-1* sg2). We have added these results in revised Figure 3f and 3g.

Revised Figure 3f-3g



10. In line 526, the author mentioned "HeLa CRISPRi cells were generated by lentiviral integration (~3 to 5 MOI) using the dCas9-KRAB-blast plasmid (Addgene #89567), followed by single cell isolation." However, this raises a question about the rationale behind the single cell isolation. It seems the procedure involves introducing a lentivirus library into the cells before isolating individual cells. If this is indeed the case, it implies the potential creation of over 795 unique cell lines, each corresponding to different single-guide RNAs (sgRNAs) for CRISPR interference (CRISPRi). Such a scenario suggests that 795 individual transfections of ribonucleoprotein complexes (RNPs) would be necessary for the screening experiment. Nevertheless, the text does not mention any details regarding the execution of these individual transfections, leaving a gap in the explanation of the methodology.

Furthermore, the methodology for screening sgRNAs and confirming their association with the signaling pathway needs to be meticulously detailed in the methods section. It is equally important that a comprehensive description of these processes is presented in the results section to ensure clarity and understanding.

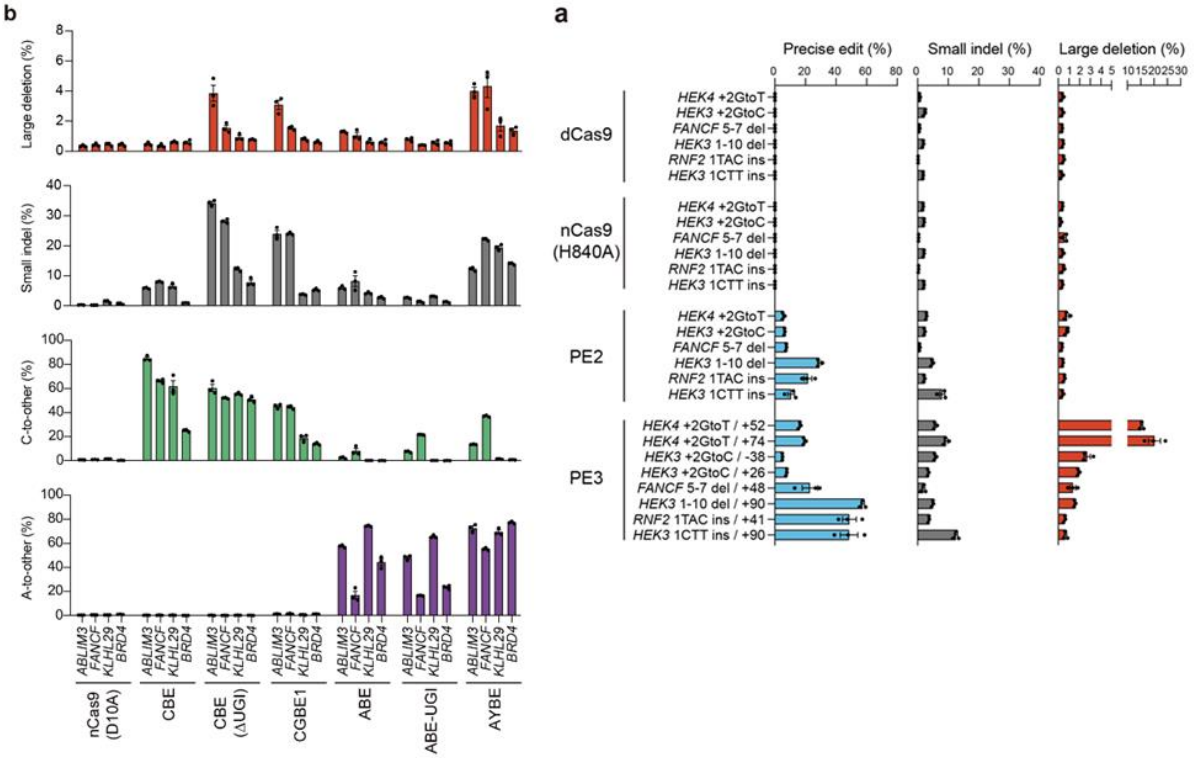
Response: We apology for the ambiguous description, potentially leading to misunderstandings. In fact, single cell was isolated to establish CRISPRi HeLa single cell line, after lenti-viral transduction with the CRISPRi library. To clarify this, we added a dedicated section to clarify the exact protocol for CRISPRi HeLa cell line construction in [lines 594-598].

11. It is important to note that both base editors and prime editors utilize nCas9 as the underlying structure. In this manuscript, the base editors mentioned employ D10A nCas9 as the base architecture, while PE2 and PE3 use H840A nCas9. Considering this perspective, it is recommended to include a control group with nCas9 in Figure 4 and Figure 5, targeting different sites.

Response: Thank you for valuable comments here. To address the reviewer's concern, we conducted all experiments with nCas9(D10A) as a BE control and with nCas9(H840A) as a

PE control. The updated data with nCas9(D10A) and nCas9(H840A) were added to revised Figure 4b-4d and 5a-5b, respectively.

Revised Figure 4b (left) and 5a (right)



12. In Figure 4, it is essential to provide a detailed version and plasmid information of the cytosine base editors (CBE), cytosine-guanine base editors (CGBE), adenine base editors (ABE), and adenine-to-cytosine base editors (AYBE).

Response: We have added all specific version of BEs in main text (lines 269-286) and plasmid information in the method section (lines 658-662).

13. In Figure 4e, the author presents a hypothesis concerning the occurrence of small indels or large deletions. However, it is important to note that similar hypothesis or part of the hypothesis has already been proposed in previous reports (PMID: 32690971; PMID: 27096365). Thus, the hypothesis lacks novelty.

Response: As the reviewer commented, the two papers suggest similar model with us. However, as the reviewer mentioned in above Q2, some studies did not see significant large deletions in the BE system [PMID: 32711379], whereas some observed BE-mediated small indels or large deletions [PMID: 32690971; PMID: 27096365], which is controversial. In our study, we concluded that BEs are reasonable for large deletion generation, especially for BE platforms that easily produce AP sites [i.e., CBE(ΔUGI), CGBE1, AYBE]. Based on our observation, we clarify BE-mediated DSBs and large deletion with a Schematic.

14. As mentioned earlier, there have been reports suggesting that base editing does not result in large deletions. It is commendable that the author has identified instances of large deletions occurring in base editors. However, the conflicting viewpoints need to be carefully addressed and discussed. To strengthen the argument, more compelling evidence and better discussion should be provided.

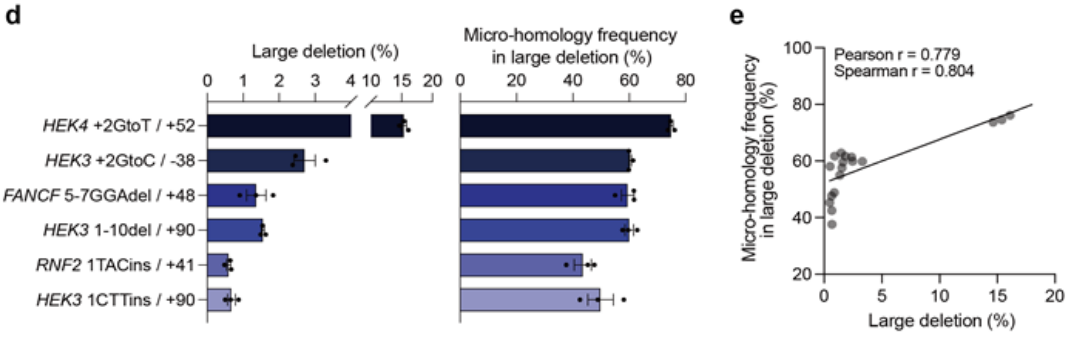
Response: To the best of our knowledge, no one have tested with various BE platforms to investigate the occurrence of DNA large deletion events. In the present study, we used 6 representing BE forms and found that BE platforms that easily produce AP sites [i.e., CBE(ΔUGI), CGBE1, AYBE] have tendencies to generate large deletion more, for the first time. KO cell line test in Figure 4e and ART treatment experiment in Figure 4f further enhance our findings.

15. In Figure 5C, we observe that the efficiency of large deletions using PE3 is categorized into two distinct groups: one group exhibits an efficiency of approximately 3%, while the other demonstrates a significantly higher efficiency around 50%. This substantial variance warrants further discussion to understand the underlying causes.

Moreover, it is important to explore the potential relationship between these higher editing efficiency and the large deletion efforts of PE2/PE3.

Response: Thank you for the comments. In fact, it is not easy to determine why PE3 generates different outcomes of small indels and large deletions in different target sites. Regarding the reviewer's comment, we speculated that the sequence context including microhomology information might be one reason. Thus, we investigated the microhomology (2-8bp) frequencies within large deletion patterns and compared the large deletion frequencies with microhomology frequencies. We observed a positive correlation between them. Therefore, we assumed that the target dependency could be one reason for generating large deletion events. We have added these results to revised Supplementary Fig. 12.

Revised Supplementary Figure 12d-12e.

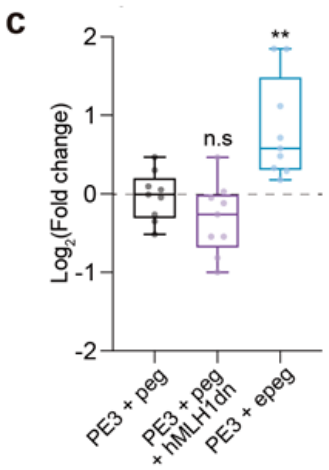


16. The results of PE3+epeg, as shown in Figure 5d, exhibit slightly significant variability. To obtain more reliable data, increasing the sample size is recommended.

Response: As the reviewer suggested, we conducted the experiments with additional epegRNAs (total 3 independent targets) and replicates (n=3). To examine the effect of

hMLH1dn and epegRNA, we displayed the data with fold change values of large deletion frequency comparing to the frequency of PE3+peg approach. We have updated data in revised Figure 5c.

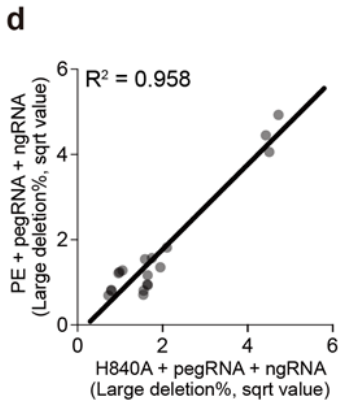
Revised Figure 5c



17. In Fig 5 and discussion section, further discussion is warranted regarding the underlying mechanism by which the PE3 system leads to a higher incidence of indels and large deletions. Previous studies have demonstrated that efficient targeting effects can be achieved through the combined action of two nCas9 enzymes (PMID: 23992846). As the PE3 system employs a dual sgRNA strategy, which shares similarities with the principle of gene knockout using nickase Cas9, it is crucial to include a control group in Figure 5. Specifically, adding nCas9 + sgRNA1 + sgRNA2 (where sgRNA1/2 corresponds to the sgRNAs utilized in PE3) might help ascertain whether the observed large indels are solely attributable to nCas9.

Response: Thank you for valuable comments here. As the reviewer suggested, we conducted nCas9 (H840A) with pegRNA and ngRNA. When we treated nCas9 (H840A) with pegRNA and ngRNA, instead of PE2 with pegRNA and ngRNA (i.e., without RT) in human cells, we found a strong correlation of large deletion frequencies between them, indicating that the generation of double nicks by nCas9 is the most critical factor for generating DSBs and large deletions in PE3. We have added this data in revised Figure 5d.

Revised Figure 5d.

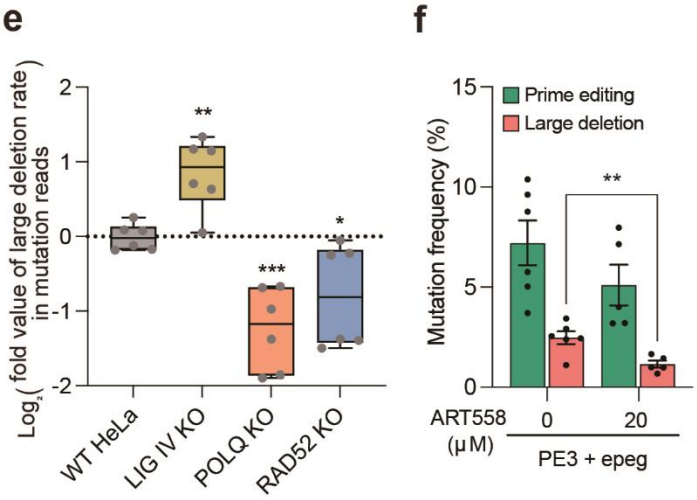


18. In Figure 5, it is evident that the authors have not provided mechanistic explanations or a hypothesis. This lack of information creates a disconnect from the "detailed mechanism" (in title) concerning prime editors. Therefore, it is crucial to design and incorporate additional experiments pertaining to the mechanism in Figure 5.

Response: As the reviewer mentioned, we generally toned down the term "detailed mechanism" overall and changed the title accordingly. However, we did our best to reveal the mechanism of PE-mediated indels and large deletion, as follows.

We further investigated subsequent pathways leading to large deletions with PE3 platform in various KO cell lines (*LIG IV* KO, *POLQ* KO, *RAD52* KO HeLa cell line) used in Figure 3e. Similar to Cas9 nucleases and BEs, the frequency of large deletions was increased in *LIG IV* KO cell line, but decreased in *POLQ* KO cell line (**Fig. 5e**). Also, the treatment of ART558 decreased the large deletion events (**Fig. 5f**), suggesting that TMEJ remains the major pathway for PE-mediated large deletions. However, we also observed that the frequency of large deletions was decreased in *RAD52* KO cell line, opposite to Cas9 nucleases (**Fig. 5e**), possibly implying the difference between DNA repair processes after Cas9- and PE-driven DSBs.

Revised Figure 5e-5f.



Reviewer #2 (Report for the authors (Required)):

The manuscript "Detailed mechanisms for unintended large DNA deletions with CRISPR, base editors, and prime editors," by Hwang, Lee, and colleagues sought to characterize the frequencies of and mechanisms by which large DNA deletion byproducts are generated during the course of genome editing experiments. To address the quantification problem, they developed a long-range next-generation sequencing (NGS) pipeline to investigate the genotypes that arise from Cas9 nuclease cutting, base editing, and prime editing. To investigate the mechanism by which these genotypes occur, they performed a CRISPRi screen and found that long deletion genotypes arise from the theta-mediated end-joining (TMEJ) pathway. Lastly, they examined a variety of base editing architectures and prime editing systems to investigate the frequencies of long deletions in those systems.

The work addresses an important question in the application of therapeutic genome editing: "How can one quantify and reduce undesired genotypes that cannot be sequenced directly with traditional NGS techniques?" However, while this question is important, there are several aspects of this work that must be addressed before it is suitable for publication in Nature Biomedical Engineering. Addressing the following points would significantly strengthen the manuscript, and improve its relevance to a biomedical engineering audience:

Response: We appreciate this reviewer's supportive and positive comments.

Major points:

- (Lines 38–39): The maximum frequencies of the deletions that the authors found for biomedically relevant base editors (CBE: 0.8%, ABE: 1.42%) are quite small. Moreover, the frequencies of these deletions are low for prime editors as well, with a maximum large-deletion frequency of 1.2% for PE2 and a clear bimodal distribution for PE3, with the majority of edits having a large-deletion frequency under 5%. The abstract as written makes a misleading and false equivalence between Cas9-based and nCas9-based technologies as generating similar levels of such byproducts. In reality, the main finding, that BE and PE result in far lower frequencies of these large deletions, is less sensationalist but much more important and accurate. The authors should reword this sentence to make this distinction more clear or remove this sentence entirely from the abstract, as it is misleading and sensationalist.

Response: We appreciate the reviewer's comment. As suggested, we have revised the abstract with the expression "fewer but significant large deletions compared to Cas nucleases" to avoid the misunderstanding in [lines 40-41].

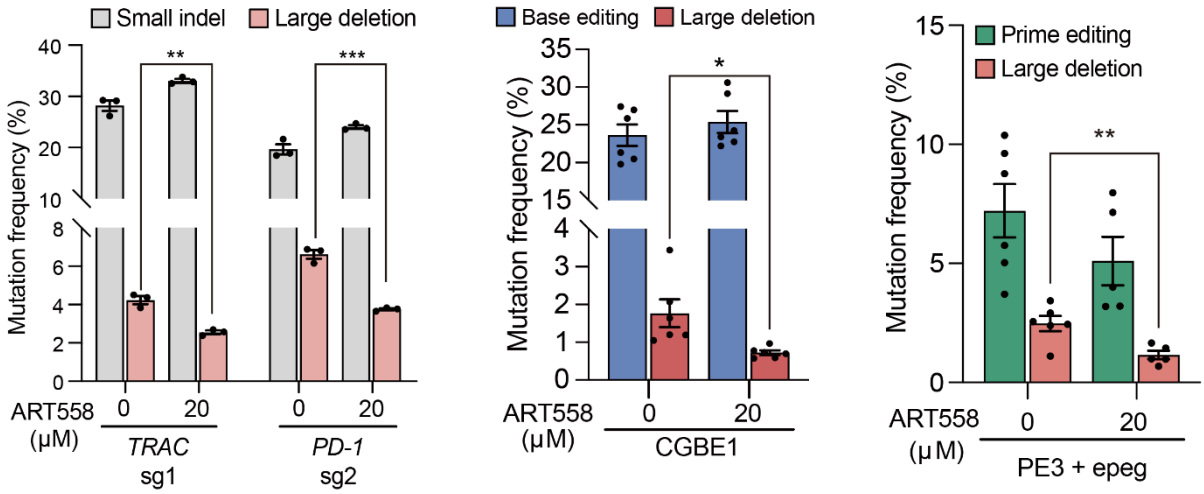
- (232–234): While M4344 works for reducing TMEJ-mediated large deletions, inhibition of ATR kinase would also preclude the HDR mechanism and would therefore be unsuitable for approaches used in many applications of Cas9 nuclease, such as integration of a chimeric antigen receptor into T cells. The authors should make this caveat more clear. Alternatively, a more convincing argument could be made using an inhibitor of DNA polymerase theta to selectively inhibit TMEJ but not HDR.

Response: Thank you for valuable comments here. To address the reviewer's concern, in

addition to M4344, we applied POLQ inhibitor (ART558) to investigate whether the generation of undesired large deletion was mitigated in HeLa cell line. We found that, similar to M4344, ART558 reduced undesired Cas9-mediated large deletion as well (added to revised Fig. 3g). As the reviewer mentioned, because the M4344 can affect the entire end resection process, the ART558 may be more suitable as a large deletion inhibitor.

In addition, we also applied ART558 to base editor (CGBE1) and prime editor (PE3). Results exhibited that ART558 effectively mitigated the undesired large deletion in both CGBE1 (added to revised Fig. 4f) and PE3 (added to revised Fig. 5f). Taken together, we suggest that ART558 can be a suitable candidate to mitigate undesired large deletions in therapeutic applications with Cas9, base editor and prime editor. We have added these results in revised Figure 3g, Figure 4f and Figure 5f, respectively.

Added Figure 3g, 4f, and 5f.



- (249–253): Importantly, CBE(ΔUGI), CGBE, and AYBE are not relevant to clinical applications of base editing—they are explicitly avoided for clinical use because their mechanism yields byproducts including deletions (though still fewer than nucleases). A better way of structuring this section is to present the canonical CBE/ABE results (which show low deletion rates), and then describe the additional non-clinical constructs tested as part of a mechanistic analysis showing how these low-frequency deletions arise. Then, the authors can add that based on their analysis, ABE–UGI could be used in cases where there are TC motifs present in the base editing window to further drive down indels resulting from bystander cytidine deamination.

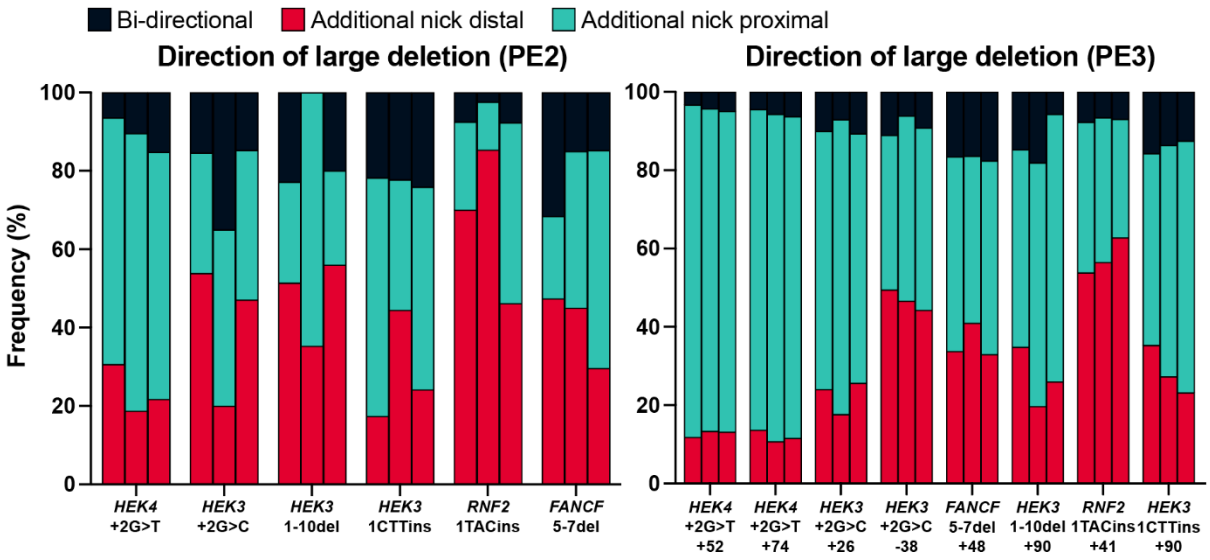
Response: Thank you for the comments. We agree with the comments that the canonical CBE or ABE have much chance to be used for further clinical applications because of low byproducts compared to CBE(ΔUGI), CGBE, and AYBE. However, in terms of expanding the knowledge of BE platforms, we believe that various BE platforms including CBE(ΔUGI), CGBE, and AYBE should be compared in parallel. Moreover, someone may improve the CGBE or AYBE as a safety form in the future. Also, we found an effort to utilize GBE (Glycosylase base editor) as a gene editing therapeutic application named *BRL-103* (ClinicalTrials.gov identifier; NCT05442346). Please understand the current structure.

- (293–294): The authors should examine the exact deletion genotypes that arise from PE3. Do these deletions extend past the two nick sites? If so, this would suggest a mechanism that also involves resection mediated by ATR kinase and therefore treatment with M4344 should reduce the frequency of these deletion genotypes. If not, then the authors should note that the mechanism by which these genotypes occur is distinct from that discussed earlier in the paper, and perhaps propose a different mechanism that can explain the results observed.

Response: Thank you for the comments.

To address the reviewer's concern, we analyzed the direction of large deletion events generated by PE2 or PE3. Basically, the origin of large deletion would be the nick site by pegRNA. The directions can be divided into three groups (Bi-directional, additional nick distal and additional nick proximal). We determined directionality when one side is twice longer than the other, and judged proximity based on the additional nick site. In case of PE3, the additional nick proximal large deletion patterns are increased compared to the PE2. We have added the data to revised Supplementary Figure 12.

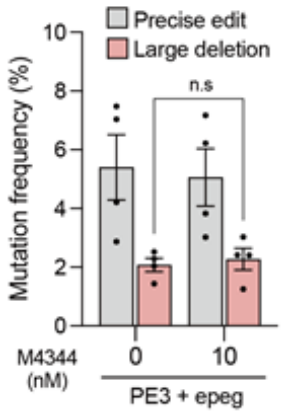
Added Supplementary Figure 12a.



We further examined whether M4344 can reduce the large deletion frequency or not. Results showed that M4344 did not contribute to reduce the large deletion frequency in PE3 approach, in contrast to the POLQ inhibitor (ART558) that reduce the large deletion frequency well. We have added the data to revised Supplementary Figure 12.

Added Supplementary Figure 12f.

f



When we treated PE3 platform in various KO cell lines (LIG IV KO, POLQ KO, RAD52 KO HeLa cell line), we found that the frequency of large deletions was increased in LIG IV KO cell line, but decreased both in POLQ KO cell line and in RAD52 KO cell line, which is different from the case of Cas9 nucleases. Taken together, we concluded that TMEJ remains the major pathway for PE-mediated large deletions but detailed mechanisms might be different between DNA repair processes after Cas9- and PE-driven DSBs.

- (452–496): Were read pairs deduplicated before processing? This step is crucial for quantitative analyses involving DNA shearing followed by a barcoding amplification step because PCR bias during the barcoding step can result in some sequences being duplicated and therefore overrepresented.

Response: Thank you for comments. In this study, we developed optimized long-range amplicon sequencing method and applied it to most experiments without the CRISPRi screening experiment. For the optimized long-range amplicon sequencing method, there are three important improvements to mitigate PCR-bias: i) selecting an optimized DNA polymerase to minimize length bias during PCR, ii) developing a specialized k-mer alignment algorithm for our purpose, and iii) implementing strategies to eliminate background noise from false-positive large deletions, as shown in revised supplementary Figure 3. Although we did not employ Unique Molecular Identifiers (UMIs), we confirmed the accuracy of our protocol and pipeline in different ways as shown in Figure 1.

Minor points:

- (136, 486): In the main text, the authors defined the cutoff for large vs. small deletions at 100 bp. However, in the methods, the authors defined the cutoff at 50 bp. Can the authors please clarify this discrepancy?

Response: Thank you for the comment. We fixed it to 100 bp.

- (318–322): New data should not be included in the discussion section. Instead, the authors can consider making a new results sub-section to discuss mischaracterization of cell lines. It would also strengthen the overall impact of the paper if the authors can make an analysis of the likelihood of such mischaracterization in current clinically relevant genome editing workflows.

Response: Thank you for the comment. For the previous Supplementary Figure 10, we have moved it to revised Supplementary Figure 3. But for the previous Supplementary Figure 11, we decided to remove it in the revised manuscript because we think this data is somewhat redundant.

- (452–496): The authors did not specify the k-mer length they used for analyses performed in this work. Additionally, the use of the letter k for both the sequence length in the k-mer hash table and the k-means clustering hyperparameter is confusing.

Response: Thank you for comments. We have clarified the words. “k” for k-mer hash was to k_{hash} and “k” for k-means clustering was to $k_{cluster}$. We used k_{hash} for 11.

Reviewer #3 (Report for the authors (Required)):

This is a nice article in which the authors highlight the issue of long deletions (>100 bp) after gene editing. Although the long deletion phenomenon has been previously described (as authors cite), here the authors develop some improved experimental and analytic assays for long-range amplification, fragmented short-read sequencing or long-read sequencing, and edit sequence alignment, conduct a genetic screen identifying end resection as an important pathway required for long deletions, and demonstrate the risks of long deletions with certain base and prime editing approaches. This will be a useful reference for future investigators seeking to comprehensively characterize gene editing outcomes including long deletion events and is a cautionary tale that long deletions may indeed be present after various forms of gene editing if only one looks carefully.

Response: We appreciate this reviewer's supportive and positive comments.

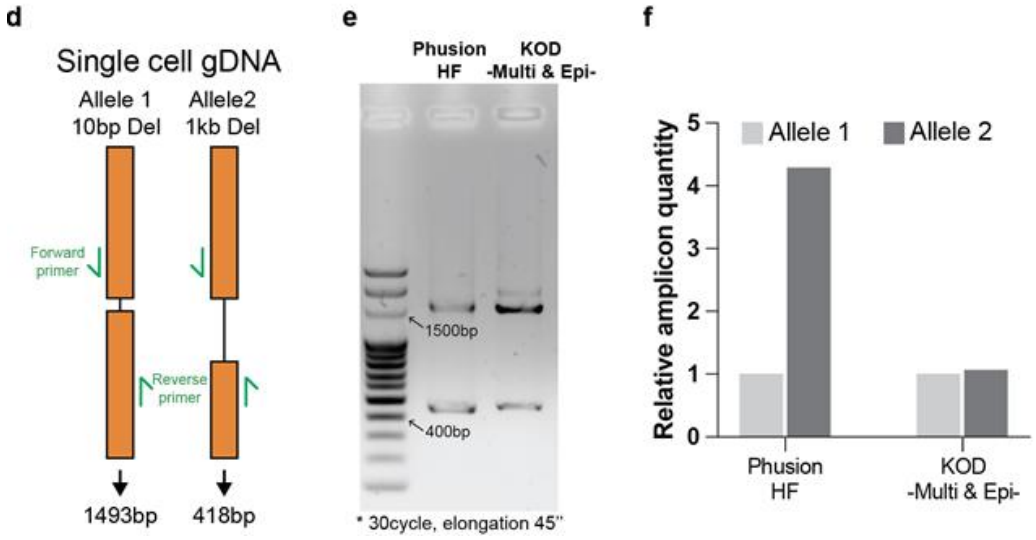
Some specific comments:

1. Fig 1c shows minimal bias comparing a ~9kb and ~10kb product (~10% difference in length). Also Supp Fig 3 is a nice design, but only tests range of 120-280 bp. What about longer deletions, e.g. up to 9 kb deletions for a 10 kb product (~90% difference in length)? It would be interesting to see the performance of the polymerases around a range of larger deletions to better understand possible length biases of the procedures.

Response: Thank you for the comments. As the reviewer pointed out, in initial experiments, we used mixtures of short fragments because it is convenient to prepare the fragments having same primer binding sequences in equimolar amounts. And we also thought that the PCR bias would be conserved between mixtures in short range from 120 bp to 280 bp (2.3-fold difference here) and mixtures in long range from 1kb to 10kb

To address the reviewer's concern, we used a single cell line in which genomic DNA has 10bp deletion on one allele and 1kb deletion on the other allele. Through PCR amplification with a same primer set, 1493bp size of PCR product is amplified from 10bp-deletion-allele and 418bp size of PCR product is amplified from 1kb-deletion-allele (3.6-fold difference here). When we tested two representative DNA polymerases – i) Phusion HF and ii) KOD -Multi & Epi- DNA polymerase which were determined to have high and low length bias, respectively, we found that the KOD -Multi & Epi- DNA polymerase still showed less bias, consistent with the results with mixture from 120 bp to 280 bp. We have added these results to Supplementary Figure 4.

Added Supplementary Figure 4d-4f.



2. Regarding the performance of the alignment strategies, it could be interesting to simulate reads with various deletions and compare approaches.

Response: Thank you for valuable comments here. As the reviewer suggested, we compared ExCas-Analyzer program with bwa mem that is a widely used alignment program and with CRISPResso2 that is a widely used analysis program for CRISPR. To this end, we produced the simulation data set for the long-range amplicon sequencing. ExCas-Analyzer has higher analysis speed and uses lower usage of memory than bwa and CRISPResso2. Additionally, we confirmed the reliability of ExCas-Analyzer using a simulated long-range sequencing data set. We generated simulation data containing large deletions at rates of 10%, 20%, 30% and 40% using the ART NGS simulator [PMID: 22199392] and checked how close ExCas-Analyzer came to the expected ratio. The table below shows the results, and R^2 was 0.8506, verifying the reliability of ExCas-Analyzer. Now, we are developing a web tool for the NGS data for Long-range amplicon sequencing to allow more researchers to freely use this method. And the fast analysis and low usage of memory will be a powerful advantage in developing web tool. We have added these results in revised Supplementary Fig. 6.

Added Supplementary Figure 6e.

Expected ratio	Proposed ratio
10%	11.82%
20%	23.13%
30%	37.14%
40%	43.26%

Program	Run time (sec)	Maximum usage of memory (MB)
ExCas-Analyzer	2.93	17
bwa mem	10.63	90
CRISPResso2	86.91	1499

3. In Fig 2, are the small insertions templated +1 position 17 reiterations or short homonucleotide insertions as have been described as a common Cas9 repair outcome (PMC5834694) or some other type of small insertion which might reflect a different repair mechanism? The anti-correlation might suggest competing DNA repair mechanisms.

Response: As the reviewer suggested, we investigated the insertion patterns for *HPRT1, ST3GAL4,* and *WISP3* in HeLa and HEK293T cell lines. Consistent with the findings reported in the study [PMC5834694], the predominant insertion pattern observed was a 1nt homonucleotide insertion, as illustrated in the figure below. For insertions longer than 1bp, the sequence adjacent to the insertion site was often repeated.

Figure for the reviewer #3 only. Various outcomes of insertion pattern after Cas9 treated.

HeLa

HPRT1

CTTGAGCACAC	C	AGAGGGCTA	68.87
TTGAGCACAC	A	AGAGGGCTAC	6.47
CTTGAGCACAC	T	AGAGGGCTA	4.91
CCCTTGAGCACAC	AC	AGAGG	4.7
CCCTTGAGCACAC	ACAC	AGAGG	1.02

ST3GAL4

CCACGACCACA	A	CAGCGGCGG	91.15
TCCCCACGACCACA	CA	CAGCGG	1.14
CCACGGCCACA	A	CAGCGGCGG	0.39
CCACGACCACA	G	CAGCGGCGG	0.39
TCCCCACGACCACA	AA	CAGCGG	0.38

WISP3_sg1

CCAGGAGGGCA	A	ACGGGGCTT	88.9
CCAGGAGGGCA	C	ACGGGGCTT	0.85
CCAGGAGGGCA	G	ACGGGGCTT	0.64
CCAGGAGGGCA	A	ACGGGGCTC	0.53
CCAGGAGGGCA	A	GCGGGGCTT	0.46

HEK293T

HPRT1

CTTGAGCACAC	C	AGAGGGCTA	56.97
TTGAGCACAC	A	AGAGGGCTAC	4.85
CTTGAGCACAC	T	AGAGGGCTA	3.64
CTTGAGCACAC	AC	AGAGG	3.64
CTTGAGCACAC	G	AGAGGGCTA	2.42

ST3GAL4

CCACGACCAC	A	ACAGCGGCGG	87.21
CCACGACCAC	AC	ACAGCGG	1.21
CCACGACCAC	AG	CAGCGGCGG	0.76
CACGACCAC	CTG	AGCGGCGGC	0.45
CCACGACCAC	C	ACAGCGGCG	0.38

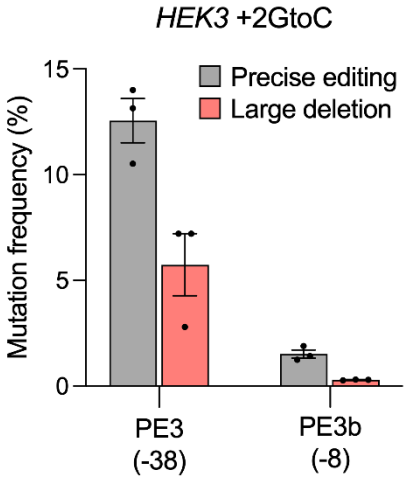
WISP3_sg1

CCAGGAGGGCA	A	ACGGGGCTT	89.88
CAGGAGGGCA	A	ACGGGGCTTC	0.58
CCAGGAGGGCA	AA	ACGGGGCTT	0.52
CCAGGAGGGCA	T	ACGGGGCTT	0.52
CCAGGAGGGCA	G	ACGGGGCTT	0.23

4. Fig 5 and Supp Fig 9: it could be interesting to try PE3b strategy to see if it is more similar to PE2 or PE3 in terms of long deletions.

Response: Thank you for the comments. As the reviewer suggested, we compared PE3 and PE3b platforms with *HEK3* +2GtoC epegRNA. Results showed that large deletion frequency of PE3b was quite decreased compared to the value of PE3. However, the prime editing frequency was also decreased because of different additional nicking gRNA. Theoretically, the PE3b may reduce the frequency of large deletions since the double nicks did not occur simultaneously, consistent with the decreased occurrence of small indels with PE3b. But, we have very few cases, we have just added description about it in the discussion section [lines 385-387].

Figure for the reviewer #3 only. Comparison between PE3 and PE3b.



5. Could the authors add some discussion regarding to what degree including a UMI might increase the accuracy and minimize the bias of the long read quantification?

Response: Thank you for the comments. In this study, we developed optimized long-range amplicon sequencing method and applied it to most experiments without the CRISPRi screening experiment. For the optimized long-range amplicon sequencing method, there are three important improvements to mitigate PCR-bias: i) selecting an optimized DNA polymerase to minimize length bias during PCR, ii) developing a specialized k-mer alignment algorithm for our purpose, and iii) implementing strategies to eliminate background noise from false-positive large deletions, as shown in revised supplementary Figure 11. Although we did not employ Unique Molecular Identifiers (UMIs), we confirmed the accuracy of our protocol and pipeline in different ways as shown in Figure 1.

However, as the reviewer suggested, long-read sequencing platform with a UMI may aid to detect large deletion. We have added description about it in the discussion section [lines 368-370].

6. A few typos:

- Supp Fig1b "bluw" typo.

- Typo line 261-2 "In Figure 4c, each large deletion frequency was divided by the average large deletion frequency with."

- Fig S10 typo "Stard"

- For lines 321 and 322, is it meant "homozygous" and "heterozygous" rather than "homologous" and "heterologous"?

Thank you for the comments. We have fixed all in the revised manuscript.