THE EMBO JOURNAL

Break-induced replication plays a prominent role in long-range repeat-mediated deletion

Qing Hu, Hongyan Lu, Hongjun Wang, Shibo Li, Lan Truong, Jun Li, Shuo Liu, Rong Xiang and Xiaohua Wu.

Review timeline:

Submission date:	10th February 2019
Editorial Decision:	20th March 2019
Revision received:	14th July 2019
Editorial Decision:	28th August 2019
Revision received:	7th September 2019
Accepted:	11th September 2019

Editor: Hartmut Vodermaier

Transaction Report:

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

1st Editorial Decision 20th March 2019

Thank you for submitting your manuscript on BIR-over-SSA preference in human repeat-mediated deletions to The EMBO Journal. We have now received comments from three expert referees, copied below for your information. As you will see, all referees consider your findings and conclusions interesting and potentially important, but at the same time raise a number of substantive concerns with the experimental evidence in support of these conclusions. In this respect, recurrent key issues are that the provided data are not sufficiently decisive to exclude various possible alternative explanations/scenarios, and that the underlying mechanisms and the roles of implicated factors (ATM, H2AX, MSH2) remain incompletely analyzed and understood.

Should you be able to adequately address these major concerns (as well as the various more specific technical and presentational points, see especially ref 3 point 1), I would be interested in considering a revised manuscript further for eventual EMBO Journal publication. In this respect, I should however point out that our single-major-revision-round policy would make it important to comprehensively respond to all points raised at this stage. While I would not insist on fully elucidating the exact roles of ATM, H2AX or MSH2 in BIR-mediated RMD, I feel that it will be essential to strengthen the basic evidence for BIR-over-SSA, e.g. by including additional SSA-specific perturbations (see ref 1), demonstrating repair synthesis and ruling out alternative POLD3 roles (refs 2&3), addressing potential complications by end-joining mechanisms (refs 2&3), and testing the situation also in non-proliferating cells (ref 1).

REFERE REPORTS

Referee #1:

Qing Hu et al have investigated the mechanism of repeat-mediated deletion (RMD). Here they

provide evidence that, in contrast to yeast, single-strand annealing is inefficient in mammalian cells when a DSB is introduced in the context of repeats. Instead, they show that Break-Induced Replication (BIR) is predominantly utilized when the DSB is close to one repeat. They present evidence supporting oncogene expression promotes BIR but not SSA, and likely reflects single-ended DSBS generated during replication stress.
They also compare trans-acting factors involved in these distinct mechanisms of RMD, and present evidence that ATM and H2AX is required for long-range BIR-mediated RMD. Further they show a partial dependency for BIR-mediated RMD, occurring within 1% and 3% divergent sequence repeats on MSH2.

General comments:
The manuscript is well presented and the data are clear. This manuscript contains a number of potentially important observations. Specifically, it suggests that SSA is inefficient in mammalian cells compared to yeast. Further it suggests that BIR mediates RMD in regions close to a repeat. Further the authors have gone on to demonstrate roles for ATM, H2AX and MSH2 in modulating the efficiency of BIR. However, the analysis of the role of these factors is somewhat cursory, and does not provide any real mechanistic insight at present.

They demonstrate that Cyclin E overexpression promotes BIR-mediated RMD, which is an important point. However, it will also be important to test what happens to these ratios when DSBs are introduced when cells are in a quiescent state. This, it might be predicted should promote SSA-mediated RMD due to the fact that these cells are not replicating. This, if true, would also be an important observation.

Throughout the manuscript the authors attempt to distinguish between SSA and BIR mechanisms of RMD. While Depletion of POLD3 is informative for BIR, RAD52 deletion leads to loss of both outcomes, and so isn't particularly helpful to distinguish mechanism. Can the authors try to target SSA specifically? For example, Figure 2D shows that if they put breaks near repeats on both sides then this is repaired independently of POLD3. However, this does not confirm this is repaired by SSA. This should be confirmed if at all possible using SSA-specific gene knockdowns. Similarly, Figures 4 E and F would benefit from a SSA-specific marker which is not required for BIR to confirm that the L268/R300 is due to SSA.

In Figure 1 E there is nearly 50% RMD still occurring in the absence of either shPOLD3 and RAD52. So what is this? Is this inefficient knockdown or is this a distinct mechanism? This issue should be addressed.

A schematic of the EGFP-MMEJ reporter should be provided in supplemental data so that readers don't have to find Truong et al 2013.

A question arises as to whether these observations are cell line specific. The authors should again point out clearly which cell lines are being used for each experiment to clarify this point. They do this in the later figures but not the earlier ones.

Figure 4B, C and D. The authors should indicate why two breaks are necessary for efficient RMD to be observed in the RMD-GFP reporter context. Surely if this was BIR then you wouldn't need two breaks to facilitate this mechanism? Again this appear to be only partially dependent on siPOLD3 (Figure 4C).

Figure 5: The authors show that ATM is required for RMD in different contexts. However, there is no indication as to how ATM facilitates either BIR or SSA. Does ATM depletion block the common step of resection? Does this block RAD52 or RPA loading?

Figure 5: Similarly, how is H2AX modulating the proposed BIR or SSA repair mechanisms? Is this also affecting resection? How does this relate to the ATM result? These studies are currently rather anecdotal and lack mechanistic insights.

Figure 6: provides interesting insights into RMD when DSBs are induced at considerable distances away. Conclusions are however drawn as to the role of H2AX at DSBS when one repeat is far away compared to a situation when both ends are relatively near. However, to make this conclusion we

are invited to compare Figure 6C with Figure 5D. However, while this might be the case, it is difficult to make this comparison as the former uses mouse embryonic stem cells while the latter uses human U2OS cells. Therefore, there could be any number of factors which contribute to the differences observed here between these distinct cell types. So this should be done in the same cell type if a comparison is to be drawn.

Minor comments:
While generally well written, there are minor corrections required with the English throughout the manuscript. For example, the opening sentence of the results might better read: SSA is believed to be a major pathway mediating recombination between two repeats to generate RMD. Etc.

Referee #2:

Understanding the molecular mechanisms responsible for repair of double-strand DNA breaks in eukaryotes is an important and active area of research. The goal of the manuscript by Hu et al. is to characterize the molecular mechanisms responsible for repeat-mediated deletions initiated by DNA breakage in mammalian cells. In particular, based on the results of this study Hu et al propose that break-induced replication (BIR) is used predominantly over single strand annealing (SSA) in mammalian cells, especially when repeats are far apart. They observed that when the distance between repeats is increased to a 1-2kb range, SSA becomes inefficient. At the same time, the authors report that BIR can act in mammalian cells over a long distance when the DSB is close to one repeat. Importantly, the authors show that BIR can repair both double-ended DSBs as well as one-ended DSBs formed at collapsed replication forks. The authors also show that successful BIR requires ATM, whereas H2AX is required for long-range BIR while suppressing short-range SSA. All in all, this study represents one of the first attempts to investigate the molecular mechanism of BIR in mammalian cells in a systematic way, along with investigations into the competition between BIR and SSA in mammals. Due to the high importance that BIR plays in genetic instability leading to cancer, the results of this research will be of great interest for the diverse readership of EMBO journal including researchers interested in the mechanisms of DNA repair, genomic rearrangements, DNA replication, maintenance of genomic integrity and cancer development. However, the authors need to respond to the following concerns/comments.

Major concerns:

1. The first concern is that BIR is a pathway that includes extensive repair synthesis, which is the main distinguishing feature of BIR responsible for all genetic instabilities associated with BIR. There is no proof of extensive synthesis or of any repair synthesis in association with BIR/RMD pathway in the paper. How can the authors exclude a crossover or half-crossover? How can it be excluded that POLD3 can suppress these pathways as well?

2. It is unclear why SSA is so much suppressed in mammalian cells when a very modest amount of resection is needed for SSA? Is it possible that this results due to competition with NHEJ or MMEJ? It will be useful to remove these pathways (for example by eliminating NHEJ by deleting Ku or by eliminating MMEJ by deleting or suppressing LIG3, Pol theta, or another involved gene) to test this idea.

3. Fig. 1E left and right shows the level of RMD for the same break position (2024/11). Why the efficiency on the left is 2-3 times higher as compared to the right? This is concerning since the level of RMD in a presence of shPOLD3 on the left seems higher than the efficiency of RMD in normal cells (when no genes are suppressed) on the right.

4. It remains unclear how oncogene expression could stimulate BIR/RMD in a system where DSB results not from oncogene overexpression, but induced by a site-specific DSB (Fig. 3A).

Minor concerns:

5. P. 10: " RMD by L268/R300 is dependent on RAD52 but independent of POLD3, suggesting that

SSA, but not BIR, is used". Is it possible that both pathways can operate in this case in the presence of POLD3?

6. P. 11: "H2AX suppresses short-range RMD mediated by both SSA and BIR. However, in mouse cells ..with repeats located 0.4Mb from each other, H2AX does not show a suppression effect on BIR/RMD while RAD52 is required. A likely interpretation is that H2AX is required for activating the repeat for recombination when the repeat is situated far away from DSB, and this requirement compensates for the suppression effect of H2AX on SSA/RMD and BIR/RMD". Instead of two opposite effects compensating each other, it is possible that H2AX simply has no effect on the long-range RMD system. See also Fig. 5D: it is unclear why there was no effect of siH2AX? Because of no role of H2AX ?

7. It will be nice to test the effect of Rad51 on BIR/RMD.

8. For the experiments involving shRNA and siRNA, it will be important to include the number of technical and biological repeats of experiment in the "Materials and Methods" section.

Referee #3:

The manuscript by Hu et al adds some potentially interesting interpretations of mechanisms of RMD in the human genome. In particular, they highly emphasize a role for BIR as a specific mechanism that may dominate in RMD. There are, however, several aspects of the manuscript writing that make overall interpretation of their findings difficult and perhaps more care taken in explaining their interpretations.
1. The first issue is with the description of the experimental design. This includes:
a. In the RMD reporter assays, were multiple independent clones used, or were the transfections just repeated on a single clone in each case.
b. Statistical analysis and significance did not seem to be indicated for many of the experiments.
c. In the experiments on shRNA, were the infections placed under selection? Transient or long-term? Multiplicity of infection?
d. What evidence is there that the various guide RNAs all work approximately equally? When multiple guide RNAs are used in the same experiment is it possible that there are a mixture of single cleavages with double cleavages and would these potentially alter the interpretations?
2. The authors are sometimes very monochromatic and definitive about their interpretations. For instance, from the beginning is the assumption that the 'single-ended' breaks are caused but cutting near one repeat and further from the other repeat. They assume this is purely a factor in excision without considering other arguments. For instance, is it possible that excision exposes both repeats but that the longer flaps on the repeats inhibit resolution of the intermediates in some way? The authors argue later on that this is likely the case for BIR when the break moves a little too far away from the repeat, but do not consider this possibility for SSA or other mechanisms, such as in-register MMEJ, as hypothesized from the Jasin lab for some of these events.
3. The authors use the excision argument and PolD3 dependence largely to argue for BIR. Although I think their argument is reasonable, I am not convinced that PolD3 can only affect BIR. The Constantino reference shows that it can influence BIR, but uses a vector system specific for that.
4. In the replication fork collapse experiments is the effect dependent on replication direction where the reporter gene is located? Wouldn't this cause different influences in different clones potentially?
5. It seems too strong to say that ATM is 'indispensable' for SSA and BIR/RMD. In fact, KO of ATM provides a relatively modest affect, while siRNA has a bigger effect. Any explanation for this anomaly in Figure 5?
6. The biggest take home message from the manuscript seems to be that BIR is the dominant mechanism for of RMD, particularly relative to SSA. The importance of BIR here is consistent with their data and actually consistent with the presence of RM duplications as well as deletions. However, it seems to be put too forcefully given that there are many thousands of bases for breaks to occur that might contribute to SSA and other mechanisms and only relatively small targets that will strongly emphasize BIR. The Stark lab showed about a 5-fold influence of cleavage very near the repeat, compared to up to 20 in this study. However, the larger target area for larger breaks could still make them a major, if not the major, contributor.

Reviewer #1:
1. They demonstrate that Cyclin E overexpression promotes BIR-mediated RMD, which is an important point. However, it will also be important to test what happens to these ratios when DSBs are introduced when cells are in a quiescent state. This, it might be predicted should promote SSA-mediated RMD due to the fact that these cells are not replicating. This, if true, would also be an important observation.

Response: We overexpressed p27 to enrich cells at the quiescent state (G0/G1) (SI Fig 6) and found that both BIR/RMD and SSA/RMD are reduced (new data in Fig 3F). This is likely due to the inhibition of end resection in G0/G1 cells, which is required for both BIR/RMD and SSA/RMD. We induced cyclin E overexpression in p27-expressing cells. While cyclin E overexpression stimulates BIR/RMD after Cas9 cleavage at R29 in cycling cells, this effect was not observed when p27 is overexpressed (Fig 3F).

2. Throughout the manuscript the authors attempt to distinguish between SSA and BIR mechanisms of RMD. While Depletion of POLD3 is informative for BIR, RAD52 deletion leads to loss of both outcomes, and so isn't particularly helpful to distinguish mechanism. Can the authors try to target SSA specifically? For example, Figure 2D shows that if they put breaks near repeats on both sides then this is repaired independently of POLD3. However, this does not confirm this is repaired by SSA. This should be confirmed if at all possible using SSA-specific gene knockdowns. Similarly, Figures 4 E and F would benefit from a SSA-specific marker which is not required for BIR to confirm that the L268/R300 is due to SSA.

Response:
It has been shown previously that while RAD51 is required for HR, it suppresses SSA in both yeast and mammalian cells[1,2]. We tested the dependence of repeat-mediated deletion (RMD) on RAD51. R11 (2024/11) cleavage of the EGFP-SSA reporter, which generates a 11 bp tail to one repeat depends on RAD51, suggesting that the BIR-RMD mechanism is RAD51 dependent (new data in Fig 1G). In contrast, RMD induced by cleavage at R1011 (1024/1011), L11/R11 and L31/R29 is increased upon RAD51 depletion, which is consistent with the usage of SSA (new data in Fig 1G and SI Fig 3). Similar results were obtained in ES cells carrying the RMD-GFP reporter. We showed that RMD is dependent on RAD51 when cleavage is at L17/R28.4k (new data in Fig 4C), which is consistent with usage of BIR/RMD, but is suppressed by RAD51 upon cleavage at L268/R300 as expected when SSA is used (new data in Fig 4G). These data support the model that SSA is used when two tails of a break are similar in size.

3. In Figure 1E there is nearly 50% RMD still occurring in the absence of either shPOLD3 and RAD52. So what is this? Is this inefficient knockdown or is this a distinct mechanism? This issue should be addressed.

Response: To test whether there is a RAD52 independent pathway to mediate RMD-BIR, we knocked out (KO) *RAD52* in U2OS (EGFP-SSA) reporter cell line using CRISPR technology (SI Fig 11A). We showed that RMD-BIR is reduced in *RAD52*-KO cells to a level comparable to that in RAD52-depleted cells; there is still a substantial level of RMD-BIR (new data in SI Fig 11B comparing to Fig 1F). This suggests that RAD52 was sufficiently depleted by our shRNAs. This also indicates that a RAD52-independent pathway to mediate RMD/BIR likely exists. POLD3 is an essential gene in mammalian cells[3], so we cannot use the same KO strategy to address this point. Currently, we are testing whether RAD52 and POLD3 are epistatic to each other, which may help define different mechanisms and pathways involved in BIR/RMD. However, this study may take substantial more time and we hope to address this point in the future publication.

4. A schematic of the EGFP-MMEJ reporter should be provided in supplemental data so that readers don't have to find Truong et al 2013.

Response: We added a schematic drawing of EGFP-MMEJ reporter in SI Fig 4.

5. A question arises as to whether these observations are cell line specific. The authors should again point out clearly which cell lines are being used for each experiment to clarify this point. They do this in the later figures but not the earlier ones.

Response: We have indicated the cell lines that are used in all the figures.

6. Figure 4B, C and D. The authors should indicate why two breaks are necessary for efficient RMD to be observed in the RMD-GFP reporter context. Surely if this was BIR then you wouldn't need two breaks to facilitate this mechanism? Again this appear to be only partially dependent on siPOLD3 (Figure 4C).
Response: We added in new data to show BIR/RMD works efficiently over a long distance (when ~ 200 kb away, still 2-3%, new data in Fig 4E). However, when the distance is increased to ~0.4 Mb range, the efficiency becomes very low (e.g. ~0.2-0.3% when ~400 kb away, L17/R400k, Fig 4E and 6A). We propose that domain activation by DSBs possibly through □H2AX is needed for RMD. In the RMD-GFP reporter, the repeats are situated far-apart (~400 kb). If only one DSB is made close to one repeat, the repeat close to the DSB is activated for recombination but the other repeat is not activated since it is too far away from the DSB and thus RMD is not efficient. We have discussed this "domain activation" mechanism in the section "A DSB generated in the vicinity of a remote inaccessible repeat stimulates BIR/RMD" (Fig 6). In most experiments, we used two breaks in order to achieve efficient RMD.

7. Figure 5: The authors show that ATM is required for RMD in different contexts. However, there is no indication as to how ATM facilitates either BIR or SSA. Does ATM depletion block the common step of resection? Does this block RAD52 or RPA loading?

Response: ATM is involved in end resection. We previously showed that ATM phosphorylates CtIP at multiple sites and this ATM-mediated phosphorylation of CtIP is important for promoting end resection, loading RPA and activating HR[4]. We anticipate that the role of ATM in end resection is critical for both RMD/BIR and RMD/SSA. We also performed experiments and showed that both BIR/RMD and BIR/SSA depend on CtIP (new data in SI Fig 13), confirming the importance of end resection for RMD/BIR and RMD/SSA. In our previous study, we have demonstrated that ATM is important for RPA loading through modulating CtIP activity[4]. Thus, we propose that ATM promotes end resection and RPA loading to facilitate BIR/RMD and SSA/RMD. We have added discussion on this point in the manuscript.

8. Figure 5: Similarly, how is H2AX modulating the proposed BIR or SSA repair mechanisms? Is this also affecting resection? How does this relate to the ATM result? These studies are currently rather anecdotal and lack mechanistic insights.

Response: It was shown previously that H2AX is required for HR but is involved in suppression of SSA[5]. These findings suggest that H2AX, in contrast to ATM, is not required for end resection. In another note, it was shown that H2AX prevents CtIP-mediated end resection in G1 lymphocytes[6], suggesting that H2AX has a role in suppressing end resection under certain conditions. As for the role of H2AX in HR, MDC1, which directly binds to □H2AX, is required for HR[7]. Despite these observations, the exact mechanisms of how H2AX is involved in promoting HR while suppressing SSA remain elusive in the field.

Although H2AX is a substrate of ATM, its function in recombination-mediated DSB repair appears to be different from ATM. Beside H2AX, many other DNA damage proteins are phosphorylated by ATM, and CtIP is the key substrate of ATM to promote end resection as mentioned above. We believe that the role of H2AX to suppress RMD/BIR and RMD/SSA is related to its general function in suppression of SSA. We have added discussions on the difference of H2AX and ATM in regulating RMD/BIR and RMD/SSA in the manuscript.

9. Figure 6: provides interesting insights into RMD when DSBs are induced at considerable distances away. Conclusions are however drawn as to the role of H2AX at DSBS when one repeat is far away compared to a situation when both ends are relatively near. However, to make this conclusion we are invited to compare Figure 6C with Figure 5D. However, while this might be the case, it is difficult to make this comparison as the former uses mouse embryonic stem cells while the latter uses human U2OS cells. Therefore, there could be any number of factors which contribute to

the differences observed here between these distinct cell types. So this should be done in the same cell type if a comparison is to be drawn.

Response: We agree that H2AX dependence may not be the same in different cell types. We thus targeted the EGFP-SSA reporter to the ROSA locus in mouse ES cells. We showed that short-range BIR-RMD in ES cells is also suppressed by H2AX as that in U2OS cells (new data in SI Fig 10).

10. Minor comments:
While generally well written, there are minor corrections required with the English throughout the manuscript. For example, the opening sentence of the results might better read: SSA is believed to be a major pathway mediating recombination between two repeats to generate RMD. Etc.

Response: We have edited the manuscript and corrected some English errors.

Reviewer #2:
1. The first concern is that BIR is a pathway that includes extensive repair synthesis, which is the main distinguishing feature of BIR responsible for all genetic instabilities associated with BIR. There is no proof of extensive synthesis or of any repair synthesis in association with BIR/RMD pathway in the paper. How can the authors exclude a crossover or half-crossover? How can it be excluded that POLD3 can suppress these pathways as well?

Response: By using shRNAs, we showed that in addition to POLD3, RMD/BIR also requires replication proteins such as PCNA, RFC and CDC45 in a manner similar to BIR assayed in yeast[8], suggesting that RMD/BIR indeed needs extensive repair DNA synthesis (new data in SI Fig 2). Although we cannot exclude the possibility that POLD3 is involved in pathways other than BIR, this study demonstrated that RMD/BIR involves a DNA replication mechanism.

2. It is unclear why SSA is so much suppressed in mammalian cells when a very modest amount of resection is needed for SSA? Is it possible that this results due to competition with NHEJ or MMEJ? It will be useful to remove these pathways (for example by eliminating NHEJ by deleting Ku or by eliminating MMEJ by deleting or suppressing LIG3, Pol theta, or another involved gene) to test this idea.

Response: We knocked down or knockout KU70 or POLQ in U2OS cells and ES cells and found that both BIR/RMD and SSA/RMD are increased in KU70- or POLQ-deficient cells (new data in SI Fig 12), suggesting that NHEJ and MMEJ pathways are in competition with BIR/SSA and BIR/SSA. However, since both BIR/SSA and BIR/SSA are suppressed by NHEJ and MMEJ, inefficient SSA/RMD (while BIR/SSA is efficient) is not simply due to competition of RMD with NHEJ and MMEJ. One possibility causing inefficient SSA is that end resection is very slow in mammalian cells, which is estimated at ~0.2 kb/h and is 20 times slower than that in yeast[9]. This slow kinetics of end resection may limit the extent of ssDNA formation in a timely manner and thus suppress SSA of the repeats over a long distance in mammalian cells. We included this in the discussion.

3. Fig. 1E left and right shows the level of RMD for the same break position (2024/11). Why the efficiency on the left is 2-3 times higher as compared to the right? This is concerning since the level of RMD in a presence of shPOLD3 on the left seems higher than the efficiency of RMD in normal cells (when no genes are suppressed) on the right.

Response: We apologize that transfection in some experiments to deliver sgRNA/Cas9 may not be well controlled as others, causing variations between different experiments. We repeated the experiments in Fig 1E (now in 1E and 1F). We used drug (puromycin) to select sgRNA/Cas9-expressing cells to make sure that all cells that are analyzed by FACS analysis contain the sgRNA/Cas9 vector.

4. It remains unclear how oncogene expression could stimulate BIR/RMD in a system where DSB results not from oncogene overexpression, but induced by a site-specific DSB (Fig. 3A).

Response: Oncogene expression induces replication stress which activates S-phase checkpoint. We showed that BIR/RMD is stimulated by oncogene expression even after DSBs are generated by Cas9. We propose that BIR is promoted by ATR-mediated S-phase checkpoint. When we depleted ATR by shRNA, BIR is reduced before cyclin E expression and this probably is due to the involvement of ATR in end resection to promote BIR[10]. Interestingly, we also found that upon oncogene expression, increased BIR-RMD shows much stronger dependence on ATR than without oncogene expression (new data in Fig 3B). This observation supports the model that ATR checkpoint is involved in promoting BIR/RMD even after DSB formation. This data also suggests that ATR checkpoint has an additional role besides end resection in promoting BIR/RMD upon oncogene expression.

5. P. 10: " RMD by L268/R300 is dependent on RAD52 but independent of POLD3, suggesting that SSA, but not BIR, is used". Is it possible that both pathways can operate in this case in the presence of POLD3?

Response: We showed that the tail length significantly influences the BIR-RMD frequency. When the tails are more than 100 bp, BIR-RMD is significantly inhibited (Fig.2B). This is likely because non-homologous tails block strand invasion (see discussion in the manuscript). Thus, we favor the idea that SSA is dominant over BIR-RMD when cleavage is made at L268/R300. In support of this, we further showed that knockdown RAD51 leads to an increase of RMD when cleavage is made at L268/R300, supporting the conclusion that SSA is involved (new data in Fig 4G).

6. P. 11: "H2AX suppresses short-range RMD mediated by both SSA and BIR. However, in mouse cells with repeats located 0.4Mb from each other, H2AX does not show a suppression effect on BIR/RMD while RAD52 is required. A likely interpretation is that H2AX is required for activating the repeat for recombination when the repeat is situated far away from DSB, and this requirement compensates for the suppression effect of H2AX on SSA/RMD and BIR/RMD". Instead of two opposite effects compensating each other, it is possible that H2AX simply has no effect on the long-range RMD system. See also Fig. 5D: it is unclear why there was no effect of siH2AX? Because of no role of H2AX?

Response: We agree that we cannot exclude the possibility that H2AX has no effect on the long-range RMD. We have added in discussion in the manuscript. On the other side, since we observed that an additional DSB close to a remote repeat stimulates RMD in a manner dependent on H2AX (Fig 6), we favor the idea that H2AX has both SSA suppression role and domain activation role for promoting RMD over a distance (see discussion in the manuscript). We also added in new data to show that H2AX suppresses short-range RMD in mES cells (new data in SI Fig 10) in a similar manner as in U2OS cells, which confirms that H2AX has a role in suppressing RMD in mES cells when repeats are close to each other.

7. It will be nice to test the effect of Rad51 on BIR/RMD.

Response: We added in new data to examine the role of RAD51 on BIR/RMD and SSA/RMD (Fig 1G, Fig 4G and SI Fig 3). RAD51 is required for BIR/RMD but suppresses SSA/RMD.

8. For the experiments involving shRNA and siRNA, it will be important to include the number of technical and biological repeats of experiment in the "Materials and Methods" section.

Response: We added that in the "Materials and Methods" section**.**

Referee #3:

1. The first issue is with the description of the experimental design. This includes:
a. In the RMD reporter assays, were multiple independent clones used, or were the transfections just repeated on a single clone in each case.

Response: In each experiment, one single clone was used. We have multiple clones carrying SSA-EGFP reporter in both U2OS cells and ES cells. The conclusion of major findings is based on the analysis of different clones. Some of the results using different clones are shown (SI Fig 1 and SI Fig 5).

b. Statistical analysis and significance did not seem to be indicated for many of the experiments.

Response: We added in significance test in all figures.

c. In the experiments on shRNA, were the infections placed under selection? Transient or long-term? Multiplicity of infection?

Response: For shRNAs, we used lentiviral vector (LKO) and drug selection. The expression of shRNAs is stable and long term, and we used single-round infection. This description is included in M&M.

d. What evidence is there that the various guide RNAs all work approximately equally? When multiple guide RNAs are used in the same experiment is it possible that there are a mixture of single cleavages with double cleavages and would these potentially alter the interpretations?

Response: We performed "Inference of CRISPR Edits (ICE)" assay and showed that the guide RNAs that we used exhibit comparable DNA cleavage activities (SI Fig 14).

We agree that when multiple sgRNAs are used, some single cleavages may also be present. For ES (RMD-GFP) cells, two sgRNAs (one at L17) yield much higher levels of RMD than single sgRNAs (SI Fig 7 and Fig 4D), suggesting that double cleavages are efficient. Since single cut has very low RMD, the presence of some single cleavage should not change our interpretation of the results.

To further confirm our findings, for some experiments, we designed new reporters and used single sgRNAs to cut. For example, in Fig 1B left, we previously showed that SSA works much more efficiently when the nonhomologous tails are short (31bp/29bp) than long (1kb/1kb) (our original data using two gRNAs). We constructed new reporters with different length of DNA between the repeats and used the same single sgRNA to cut all the reporters. We obtained similar results as using double sgRNAs that SSA becomes inefficient when nonhomologous tails are long (Fig 1B, right).

2. The authors are sometimes very monochromatic and definitive about their interpretations. For instance, from the beginning is the assumption that the 'single-ended' breaks are caused but cutting near one repeat and further from the other repeat. They assume this is purely a factor in excision without considering other arguments. For instance, is it possible that excision exposes both repeats but that the longer flaps on the repeats inhibit resolution of the intermediates in some way? The authors argue later on that this is likely the case for BIR when the break moves a little too far away from the repeat, but do not consider this possibility for SSA or other mechanisms, such as in-register MMEJ, as hypothesized from the Jasin lab for some of these events.

Response: We agree that we should discuss more possibilities and have added in more discussions in the manuscript. We discussed the possible mechanisms to suppress BIR/RMD by nonhomologous tails. We also discussed that slow end resection in mammalian cells may be a major factor to inhibit SSA when DSBs are far away from the repeats while the size of flaps may not have significant effects. Regarding the possibility of in-register MMEJ, we performed experiments by knocking down POLQ, the key player of MMEJ and found that RMD is increased (SI Fig 12). These data suggest that in-register MMEJ is not likely used in our reporter assay systems carrying identical repeats, and MMEJ suppresses BIR/RMD and SSA/RMD.

3. The authors use the excision argument and PolD3 dependence largely to argue for BIR. Although I think their argument is reasonable, I am not convinced that PolD3 can only affect BIR. The Constantino reference shows that it can influence BIR, but uses a vector system specific for that.

Response: We agree that POLD3 may also be involved in other pathways. We added in new data to analyze the role of RAD51 in BIR/RMD and SSA/RMD. We showed that RAD51 is required for RMD when one cleavage is close to one repeat, supporting the involvement of BIR, whereas RAD51 suppresses RMD when SSA is used (Fig 1G and Fig 4C, 4G). We also showed that CDC45, PCNA and RFC, which are important in BIR shown in yeast[8], are needed for BIR/RMD (SI-Fig2). These data support our model of BIR/RMD.

4. In the replication fork collapse experiments is the effect dependent on replication direction where the reporter gene is located? Wouldn't this cause different influences in different clones potentially?

Response: We introduced the EGFP-SSA reporter randomly into the genome of U2OS cells and examined different single clones carrying the reporter. Southern blot analysis showed that these clones are single-copy integration at different sites in the genome. All the clones showed a higher level of RMD induced by a nick than a DSB (Fig 3D, right and SI Fig 5). Since the reporter is integrated randomly into the genome, these data support the idea that replication direction may not have a big impact on inducing BIR/SSA.

5. It seems too strong to say that ATM is 'indispensable' for SSA and BIR/RMD. In fact, KO of ATM provides a relatively modest affect, while siRNA has a bigger effect. Any explanation for this anomaly in Figure 5?

Response: We changed the word "indispensable" to "involved". In Fig 5B left, we used ATM inhibitor (ATMi) not siRNA. It was described previously that inhibition of ATM activity often results in more severe effect than *ATM* KO. Zha's lab showed that mice carrying homozygous kinase-dead (KD) mutations in *Atm* (*Atm* KD/KD) died during early embryonic development, while *ATM*-KO mice are viable, and their work also suggested that ATM-KD protein elicits additional inhibitory effects in HR[11]. It is believed that ATM-KD has a dominant-negative effect and thus the effect is stronger than KO. Our results that ATM inhibitor has a stronger effect is in line with these observations.

6. The biggest take home message from the manuscript seems to be that BIR is the dominant mechanism for of RMD, particularly relative to SSA. The importance of BIR here is consistent with their data and actually consistent with the presence of RM duplications as well as deletions. However, it seems to be put too forcefully given that there are many thousands of bases for breaks to occur that might contribute to SSA and other mechanisms and only relatively small targets that will strongly emphasize BIR. The Stark lab showed about a 5-fold influence of cleavage very near the repeat, compared to up to 20 in this study. However, the larger target area for larger breaks could still make them a major, if not the major, contributor.

Response: Our observations support Stark lab's findings that RMD can work over a long distance and also suggest further that BIR plays an important role to mediate the long-range RMD. When repeats are situated at different locations in the genome, one close-by cleavage at one repeat is sufficient to induce BIR/RMD while the other break can be over a long distance. We added in new data to show that BIR/RMD occurs at substantial levels (~2-3 %) even when the break is ~100-200 kb away from the second repeat (Fig 4E). This means that a break close to one repeat can find another repeat to recombine over a distance of hundreds kbs. On the other side, SSA requires two breaks that are close to both of the repeats and this requirement is much more stringent than for BIR/RMD and would significantly reduce the SSA usage compared to BIR/RMD in the genome. We agree that there might be other possible mechanisms involved in RMD, but for this study we focus on the comparison of BIR/RMD and SSA/RMD. We have added this in the discussion.

We found that the tail length drastically influences the RMD frequency. For example, when we fix one nonhomologous tail as 28.4 kb, increase of the other tail from 17 bp to 268 bp reduces RMD about 20-fold (Fig 4B). However, if we fix the small tail as 17 bp and increase the other tail from 300 bp to 28.4 kb, the reduction of RMD is about 3-4-fold (Fig 4D). In Stark's paper, they used different cleavage combination from ours (One tail is fixed as 268 bp and the other tail varies in size) and difference in RMD frequency is expected.

Reference:

1. Stark, J.M., Pierce, A.J., Oh, J., Pastink, A. & Jasin, M. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol Cell Biol* **24**, 9305-16 (2004).
2. Ivanov, E.L., Sugawara, N., Fishman-Lobell, J. & Haber, J.E. Genetic requirements for the single-strand annealing pathway of double-strand break repair in Saccharomyces cerevisiae. *Genetics* **142**, 693-704 (1996).

3. Murga, M. et al. POLD3 Is Haploinsufficient for DNA Replication in Mice. *Mol Cell* **63**, 877-83 (2016).
4. Wang, H. et al. The interaction of CtIP and Nbs1 connects CDK and ATM to regulate HR-mediated double-strand break repair. *PLoS Genet* **9**, e1003277 (2013).
5. Xie, A. et al. Control of sister chromatid recombination by histone H2AX. *Mol Cell* **16**, 1017-25 (2004).
6. Helmink, B.A. et al. H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes. *Nature* **469**, 245-9 (2011).
7. Xie, A. et al. Distinct roles of chromatin-associated proteins MDC1 and 53BP1 in mammalian double-strand break repair. *Mol Cell* **28**, 1045-57 (2007).
8. Lydeard, J.R. et al. Break-induced replication requires all essential DNA replication factors except those specific for pre-RC assembly. *Genes Dev* **24**, 1133-44 (2010).
9. Cruz-Garcia, A., Lopez-Saavedra, A. & Huertas, P. BRCA1 accelerates CtIP-mediated DNA-end resection. *Cell Rep* **9**, 451-9 (2014).
10. Jazayeri, A. et al. ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nat Cell Biol* **8**, 37-45 (2006).
11. Yamamoto, K. et al. Kinase-dead ATM protein causes genomic instability and early embryonic lethality in mice. *J Cell Biol* **198**, 305-13 (2012).

2nd Editorial Decision 28th August 2019

Thank you for submitting your revised manuscript for our consideration. With some delay related to the summer vacation period, it has now been re-evaluated by the three original referees. As you will see, the referees all consider the study significantly improved and many of the earlier concerns satisfactorily addressed. However, referees 2 and 3 still retain a number of reservations, concerning both descriptions and interpretation of data/experimental, as well as a more balanced presentation according to the present state of the literature in the field.

Given the significant amount of revision work that has already gone into this study, and the fact that the remaining issues seem to be in principle all addressable, I would be willing to give you the opportunity to address them through an (exceptional) second round of revision in this case. Please note however that this will have to be the final scientific revision here, and that it will be essential to decisively satisfy the open points, with the revisions and responses marked in the final text as well as explained in another detailed point-by-point response letter.

REFEREE REPORTS

Referee #1:

The authors have now addressed the majority of my concerns, and the manuscript is now significantly improved.

Referee #2:

The authors thoroughly addressed the reviewer's comments. They performed many additional experiments. The results of these experiments significantly improved the manuscript. They also included many additional explanations, discussions, schematics. All in all, this manuscript is significantly improved and will be of great interest for the diverse readership of EMBO journal.

The remaining minor comments:

1. Fig. 1E: Did expression of shPOLD3 affect cell viability? Is it possible that prolonged expression of shPOLD3 promoted accumulation of compensatory mutations or mechanisms?

2. Fig. 1F: The level of Induced RMD in R1011 is very low. How is it possible to detect the significant difference between such small levels of RMD in wt vs shRAD52?

3. Fig. 1G: In R11 shRAD51: is it likely that the induced RMD in this case (in the presence of shRAD51) represents in fact SSA?

4. Fig. 3A: Why the level of induced RMD is so low in the absence of DOX? This value appears 10x lower than for the same construct in Fig. 2B.

5. p.7, line 9: "...and suggests that BIR/RMD is RAD51-dependent". This is not the first example in the history suggesting that BIR is Rad51-dependent. Please cite the studies performed previously in yeast and/or other organisms demonstrating Rad51-dependence of BIR.

6. The competition between BIR and SSA has been previously described by Vaze et al, 2002. Please cite this work.

7. p.18, line 5: "invasion of the homologous sequences from the donor to the recipient..." Is it actually another way around?

8. p.18, line 15: There is no evidence for the role of Rad1/Rad10 or for ERCC1/XPF in the flap removal during BIR.

9. p.19, line 8: " Such a DSB would be repaired by other pathways such as NHEJ or HR". But..BIR is a type of HR, so not clear why HR is mentioned here as opposite to BIR.

Referee #3:

Although the authors have addressed many of the concerns of the three reviewers, there are still some difficulties with the presentation. I continue to think that the core observation looks good and the evidence that BIR is a major player in the RMDs is important. Concerns with the existing manuscript include:
1. At the bottom of page 4 the authors state that the prevalent thought is that SSA is the dominant way to mediate recombination of two repeats, resulting in RMD. That might have been true some time ago, but some of the major players in the field have been publishing otherwise for some years. The Jasin lab focuses more on the possibility of 'in-register' MMEJ and the Lupski lab has been very vocal that they feel that most Alu recombination is FOSTES/MMBIR based on their studies of naturally occurring recombination events. At least for mismatched substrates, which are the most common available, SSA is not considered a major player. Thus, this manuscript presents data supporting and quantitating factors controlling the Lupski labs thoughts.
2. In the first section of results, the SSA comment is repeated. The authors then utilize the EGFP-SSA vector for figure 1. It is not obvious to me that this 'RMD' vector is limited to SSA and that it might only be a EGFP-RMD vector. Thus, is it correct to say in the second paragraph that 'SSA/RMD frequency is much higher when the DSBs are 30 bp away ....'. Is the authors argument not that this is due to BIR? If so, the terminology needs to be more consistent, i.e. RMD if mechanism is not distinguished, SSA/RMD or BIR/RMD only if the mechanisms are clear and separable.
3. In the first full paragraph on page 6, the authors discuss their data on the influence of proximity of the break to the repeat, the authors fail to mention that this is in complete agreement with the previously published data from the Stark lab.
4. Overall there continues to be a poor description of experimental design in the results. For instance, in the first full paragraph of page 7, the authors stae 'We showed that inactivation of RAD51 results in increased RMD.....'. At no point in the paragraph is the actual experiment described. Was it shRNA or CRISPR, pools of cells following selection or individual clones. I have to go to the figure legend to see that it is shRNA and even then I don't know how the cells were handled or how the controls were handled. This type of skimpy description of experimental approaches is common in the results, making it difficult or impossible to understand the details of the experiments. Another example is on page 13, where their experimental description only includes

'we showed that H2AX inhibits both SSA/RMD...' and you need to look at the figure to see that it is shRNA and it is not clear what was done to the control line, and even after looking at the methods it is rarely clear exactly how each expt was handled in terms of cloning, pooling or selection procedures.
5. On page 9, the authors discuss oncogene expression as a form of replication stress. The use of oncogene in such a general way is unwarranted. Oncogenes refer to many pathways affecting all kinds of cell processes. These experiments refer only to one oncogene affecting cell cycle and therefore this description needs to be much more specific.

2nd Revision - authors' response 7th September 2019

Reviewer #2:
1. Fig. 1E: Did expression of shPOLD3 affect cell viability? Is it possible that prolonged expression of shPOLD3 promoted accumulation of compensatory mutations or mechanisms?

Response: We did not notice any change of cell growth after expression of shPOLD3 for the first week. However, cells expressing shPOLD3 started to grow more slowly than the control cells around 10 days after shPOLD3 expression. All the experiments were carried out immediately after shPOLD3 expression and completed 5-6 days after shPOLD3 expression. Since we analyzed the immediate effect of shPOLD3 depletion, the results are not likely due to compensatory mutations or mechanisms.

2. Fig. 1F: The level of Induced RMD in R1011 is very low. How is it possible to detect the significant difference between such small levels of RMD in wt vs shRAD52?

Response: Response: We counted $2x10^5$ cells per experiment for determining recombination frequency shown in Fig.1F. Counting this number of cells is sufficient to obtain reliable data even when the recombination frequency is down to 0.1%. Statistic analysis also showed significant difference in RMD with or without expressing shRAD52.

3. Fig. 1G: In R11 shRAD51: is it likely that the induced RMD in this case (in the presence of shRAD51) represents in fact SSA?

Response: We believe that when RAD51 is absent, BIR/RMD can also be carried out by RAD51-independent pathway but with reduced efficiency. We are using a new BIR reporter to examine RAD51-dependent and RAD51-independent BIR pathways in mammalian cells and the results will be presented in a separate manuscript in the near future. We believe BIR/RMD uses the same mechanisms as the general BIR.

4. Fig. 3A: Why the level of induced RMD is so low in the absence of DOX? This value appears 10x lower than for the same construct in Fig. 2B.

Response: We have obtained three independent EGFP-SSA (3427) cell lines with the reporter integrated in different places in the genome, and SSA frequency is different in these three cell lines (3427-17, 3427-3 and 3427-14) as shown in SI Fig 1A after I-SceI cleavage. In the most data shown in the manuscript, 3427-17 cell line was used and repair frequency was determined after gRNA cleavage. For testing the effect of cyclin E on BIR/RMD, we obtained good cyclin E inducible clones from 3427-3 cell line and this line exhibited relatively low SSA frequency compared to 3427-17 cell line. Fig. 3A used 3427-3 (tet-cyclin E) cell line, whereas Fig. 2B used 3427-17 cell line. We have indicated in the figure legends and M&M which 3427 reporter lines were used for each experiment.

5. p.7, line 9: "...and suggests that BIR/RMD is RAD51-dependent". This is not the first example in the history suggesting that BIR is Rad51-dependent. Please cite the studies performed previously in yeast and/or other organisms demonstrating Rad51-dependence of BIR.

Response: We have cited references for RAD51-dependent BIR in yeast and in other organisms.

6. The competition between BIR and SSA has been previously described by Vaze et al, 2002. Please cite this work.

Response: We have included Vase et al 2002 reference on page 6 and in the Discussion, page 16.

7. p.18, line 5: "invasion of the homologous sequences from the donor to the recipient..." Is it actually another way around?

Response: Thank the reviewer for pointing this out. We have corrected it.

8. p.18, line 15: There is no evidence for the role of Rad1/Rad10 or for ERCC1/XPF in the flap removal during BIR.

Response: As mentioned by the reviewer, the evidence of the involvement of Rad1/Rad10 and ERCC1/XPF is only in SSA and the role of these endonucleases in BIR/RMD have not been determined yet. In the Discussion, we proposed that ERCC1/XPF may be responsible for removing the flaps formed after strand invasion in BIR/RMD (Fig 1D, right, pink arrow). We have revised the relevant part in the Discussion.

9. p.19, line 8: " Such a DSB would be repaired by other pathways such as NHEJ or HR". But..BIR is a type of HR, so not clear why HR is mentioned here as opposite to BIR.

Response: We changed HR to gene conversion (GC).

Reviewer #2:
Although the authors have addressed many of the concerns of the three reviewers, there are still some difficulties with the presentation. I continue to think that the core observation looks good and the evidence that BIR is a major player in the RMDs is important. Concerns with the existing manuscript include:

1. At the bottom of page 4 the authors state that the prevalent thought is that SSA is the dominant way to mediate recombination of two repeats, resulting in RMD. That might have been true some time ago, but some of the major players in the field have been publishing otherwise for some years. The Jasin lab focuses more on the possibility of 'in-register' MMEJ and the Lupski lab has been very vocal that they feel that most Alu recombination is FOSTES/MMBIR based on their studies of naturally occurring recombination events. At least for mismatched substrates, which are the most common available, SSA is not considered a major player. Thus, this manuscript presents data supporting and quantitating factors controlling the Lupski labs thoughts.

Response: We agree that different models have been suggested to account for recombination and translocation between repetitive sequences. However, SSA is still regarded as a major pathway for the recombination between identical repeats in the field. In Jasin's pioneering work, they stated that "translocation between identical repeats predominantly arose from repair by SSA pathway (85%)" and only for diverged Alu elements, "in-register" MMEJ is used (Elliott et al., 2005). Lupski's lab has focused on recombination between *Alu* elements in the human genome which on average contain 71% sequence homology. They proposed that recombination between divergent *Alu*s is not purely mediated by homologous recombination, but by microhomology-mediated FoSTeS/MMBIR, based on the sequencing analysis data on genome breakpoint junctions. They agreed that SSA is used for the recombination between identical *Alu*s but proposed that when *Alu*s are divergent, MMBIR, but not SSA, is used for *Alu*-mediated rearrangements. One important part of their model is that replicative mechanisms are involved in rearrangements of divergent *Alu*-elements through MMBIR. Our study further support the model that replicative mechanisms are even used when repeats are identical via BIR. We have added a new section in Discussion to discuss the involvement of replicative mechanisms in mediating rearrangement of repetitive sequences and acknowledged MMBIR model proposed by Lupski's group.

2. In the first section of results, the SSA comment is repeated. The authors then utilize the EGFP-SSA vector for figure 1. It is not obvious to me that this 'RMD' vector is limited to SSA and that it might only be a EGFP-RMD vector. Thus, is it correct to say in the second paragraph that

'SSA/RMD frequency is much higher when the DSBs are 30 bp away ....'. Is the authors argument not that this is due to BIR? If so, the terminology needs to be more consistent, i.e. RMD if mechanism is not distinguished, SSA/RMD or BIR/RMD only if the mechanisms are clear and separable.

Response: In Fig 1B (left), when the two nonhomologous tails are similar in size (31/29bp, 651/634bp and 1024/1011bp), SSA (but not BIR) mechanism is used (in all three cases). This is also shown in Fig 2D that when two small tails are similar in size, SSA is the mechanism (11/11 and 31/29). BIR/RMD is only used when one tail is short but the other tail is long (e.g. 2014/11bp and 2006/29bp, Fig 2C).

3. In the first full paragraph on page 6, the authors discuss their data on the influence of proximity of the break to the repeat, the authors fail to mention that this is in complete agreement with the previously published data from the Stark lab.

Response: Thank the reviewer for raising this point. We have mentioned in the text that our results are in agreement with the previous findings from Stark's lab.

4. Overall there continues to be a poor description of experimental design in the results. For instance, in the first full paragraph of page 7, the authors state 'We showed that inactivation of RAD51 results in increased RMD.....'. At no point in the paragraph is the actual experiment described. Was it shRNA or CRISPR, pools of cells following selection or individual clones. I have to go to the figure legend to see that it is shRNA and even then I don't know how the cells were handled or how the controls were handled. This type of skimpy description of experimental approaches is common in the results, making it difficult or impossible to understand the details of the experiments. Another example is on page 13, where their experimental description only includes 'we showed that H2AX inhibits both SSA/RMD...' and you need to look at the figure to see that it is shRNA and it is not clear what was done to the control line, and even after looking at the methods it is rarely clear exactly how each expt was handled in terms of cloning, pooling or selection procedures.

Response: We have indicated shRNA depletion in the text for all relevant experiments. We also added in more description of experimental details in M&M (such as cell pooling and drug selection).

5. On page 9, the authors discuss oncogene expression as a form of replication stress. The use of oncogene in such a general way is unwarranted. Oncogenes refer to many pathways affecting all kinds of cell processes. These experiments refer only to one oncogene affecting cell cycle and therefore this description needs to be much more specific.

Response: We revised the text and made the description more specific for cyclin E.

Reference:
Elliott, B., Richardson, C., and Jasin, M. (2005). Chromosomal translocation mechanisms at intronic alu elements in mammalian cells. Mol Cell *17*, 885-894.

Accepted 11th September 2019

Thank you for submitting your final revised manuscript for our consideration. I am pleased to inform you that we have now accepted it for publication in The EMBO Journal.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Xiaohua Wu
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2019-101751R

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

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

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

2. Captions

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

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

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

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

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

B- Statistics and general methods

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

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	Our cell biology study followed the field's best practice in designing the experiments with adequate sample size and repeated at least three times independently to ensure we can detect even a small but significant effect.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	We used transfection or infection to deliver shRNAs and endonucleases (e.g. I-SceI and Cas9). When the transfection or infection efficiency is low as judged by drug resistance (<50%), we excluded the experiments from further analyses.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	Yes. We randomly assigned cells to treatment and control groups.
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	The experiments were repeated by different investigators.
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes. We used t-test to analyze the significance of the difference between samples.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes. Normality is assessed by graphical inspection and K-S test of necessary.
Is there an estimate of variation within each group of data?	Yes

USEFUL LINKS FOR COMPLETING THIS FORM

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

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

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

http://datadryad.org

http://figshare.com

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

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

http://biomodels.net/

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

Is the variance similar between the groups that are being statistically compared?	Yes+E56:G57

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	We listed catalog numbers of all antibodies used .
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	The cell lines we used were purchased from ATCC and verified the by the vendor. Cultures are maintained for no more than 20 passages, and a new culture is started from frozen cell stocks to keep the cells always at low passage number.

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

D- Animal Models

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

E- Human Subjects

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

F- Data Accessibility

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

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA