EMBO reports

The DNA Damage Response acts as a safeguard against harmful DNA-RNA hybrids of different origins

Sonia Barroso, Emilia Herrera-Moyano, Sergio Muñoz, María García-Rubio, Belén Gómez-González, Andrés Aguilera

Review timeline:

Submission date:	16 October 2018
Editorial Decision:	26 November 2018
Revision received:	26 March 2019
Editorial Decision:	19 April 2019
Revision received:	4 June 2019
Editorial Decision:	25 June 2019
Revision received:	28 June 2019
Accepted:	1 July 2019

Editor: Esther Schnapp

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 26 November 2018

Thank you for your patience while your manuscript was peer-reviewed at EMBO reports. We have finally received the full set of referee comments that is pasted below.

As you will see, all referees acknowledge that the findings are potentially interesting. However, they also all point out that the data are not sufficiently strong to support the main conclusions, and they all make several suggestions for how the study could be strengthened. I think all suggestions make sense and several also overlap, so I would like to invite you to address all referee concerns. Especially all technical concerns, such as missing controls, quantifications, statistical analyses, verifications, etc. should be addressed, and the formation of R-loops in the different phases of the cell cycle should be analyzed. I have not asked the referees for cross-comments on each others' reports because of the already delayed decision, but please let me know if you feel that any of the raised points cannot or should not be addressed, and we can discuss this further.

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

Revised manuscripts should be submitted within three months of a request for revision; they will otherwise be treated as new submissions. Please contact us if a 3-months time frame is not sufficient for the revisions so that we can discuss this further. Given the 7 main figures, I suggest that you layout the manuscript as a full article.

Supplementary figures, tables and movies can be provided as Expanded View (EV) files, and we can

offer a maximum of 5 EV figures per manuscript. EV figures are embedded in the main manuscript text and expand when clicked in the html version. Additional supplementary figures will need to be included in an Appendix file. Tables can either be provided as regular tables, as EV tables or as Datasets. Please see our guide to authors for more information.

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

We now strongly encourage the publication of original source data with the aim of making primary data more accessible and transparent to the reader. The source data will be published in a separate source data file online along with the accepted manuscript and will be linked to the relevant figure. If you would like to use this opportunity, please submit the source data (for example scans of entire gels or blots, data points of graphs in an excel sheet, additional images, etc.) of your key experiments together with the revised manuscript. Please include size markers for scans of entire gels, label the scans with figure and panel number, and send one PDF file per figure.

When submitting your revised manuscript, we will require:

- a complete author checklist, which you can download from our author guidelines (http://embor.embopress.org/authorguide#revision). Please insert page numbers in the checklist to indicate where in the manuscript the requested information can be found. The completed author checklist will also be part of the RPF (see below).
- a letter detailing your responses to the referee comments in Word format (.doc)
- a Microsoft Word file (.doc) of the revised manuscript text
- editable TIFF or EPS-formatted figure files in high resolution. In order to avoid delays later in the process, please read our figure guidelines before preparing your manuscript figures at: http://www.embopress.org/sites/default/files/EMBOPress_Figure_Guidelines_061115.pdf

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

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

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

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

REFEREE REPORTS

Referee #1:

In this study the authors performed a small-scale DDR siRNA screen to identify genes whose knockdown increased R loop levels. They found that reduction in ATR/CHK1, ATM/CHK2, and post-replicative repair proteins caused an increase of DNA-RNA hybrids and DNA damage. H3S10-P and H3K9me2 also accumulated with R loops. Increased fork asymmetry was observed after depletion of proteins in the ATR pathway. The authors proposed that the damage checkpoint and post-replicative repair are safeguards against R loop induced DNA damage. Although there are some interesting results in this manuscript, many of the experiments are not carefully controlled and interpreted. Several of the main conclusions of this are not yet convincing. This study still needs to be substantially strengthened to be suitable for publication.

1. In Fig. 1C, depletion of several DDR proteins led to accumulation of cytoplasmic R loops. This is surprising because most if not all of these DDR proteins are exclusively nuclear. Why would loss of DDR proteins cause an increase of cytoplasmic R loops?

2. In Fig. 2, the effects of DDR protein knockdown on transcription should be analyzed. If the transcription levels of these genes are altered, the R loops could be affected indirectly. If the transcription of these genes is cell cycle regulated, the effects of siRNAs on the cell cycle should also be tested too.

3. In Fig. 3, the -/+ RNH1 samples for each siRNA should be directly compared (as in Fig. 4). It does not make sense to compare different siRNAs with control siRNA in the presence of RNH1. From the data, it is impossible to tell whether RNH1 worked in control samples.

4. The data in Fig. 5 are confusing. Fork stalling is presumably the reason for fork asymmetry. If R loops increase fork asymmetry, they should affect fork rate too. It is confusing why R loops increased fork asymmetry but did not reduce fork rate.

5. In Fig. 5, the -/+ RNH1 samples for each siRNA should be directly compared. It does not make sense to compare different siRNAs with control siRNA in the presence of RNH1.

6. The data in Fig. 6 only suggest that R loops are formed in a transcription dependent manner, but they don't necessarily suggest that R loops are formed before replication in the absence of DDR proteins.

7. RAD18 has functions other than post-replicative repair. The data on siRAD18 may be over interpreted.

Referee #2:

Here the authors test the idea that DDR factors could be important for the detection of R-loops and/or to alleviate their toxicity. They carry out a targeted siRNA screen in HeLa cells to identify DDR factors whose depletion leads to the accumulation of DNA:RNA hybrids. DNA:RNA hybrids were detected using the well-characterized S9.6 antibody as well as chromatin marks that were previously shown to be associated with the accumulation of toxic DNA:RNA hybrids. The authors eventually focus their attention on a subset of 9 DDR factors that belong to different DDR pathways and conclude that pre-existing DNA:RNA hybrids could interfere with DNA replication and that DNA:RNA hybrids could also be promoted by unrepaired DSBs. Neither of these two conclusions is particularly novel. Nevertheless, the characterization of the role of DDR factors in preventing or eliminating DNA:RNA hybrids should be of interest to the wider community and could justify publication. However, there are a number of significant issues with the data that need to be addressed first. The demonstration that R-loops accumulate in the absence of some DDR factors (Figures 1,2 and 4) needs to be improved. In addition, the conclusion that depletion of 9-1-1/ATR/CHK1 causes hybrid-dependent fork stalling and DSB needs to be strengthened significantly. Finally, the authors suggest that different DDR mutants might accumulate R-loops in different phases of the cell-cycle (prior or after DNA replication). This would be a very new and very interesting result and the authors have the technical know-how to test this exciting possibility.

Major points:

1. Figures 1&2 aim to demonstrate that DNA:RNA hybrids accumulate when the expression of a number of DNA Damage Response genes is down-regulated using siRNA. This was done using S9.6 IF to give an overview of the DNA:RNA signal in the cell (Figure 1) and using DRIP to quantify DNA:RNA accumulation at two candidate loci (Figure 2). There are several important issues with these experiments:
• On Figure 1, the specificity of the S9.6 signals was not validated by a treatment with RNase H. In the end, the accumulation of DNA:RNA hybrids was only rigorously tested at two loci, RPL13A and APOE (Figure 2).
• In a number of siRNA conditions, the background signal (as determined by DRIP enrichment after

RNase H treatment) seems to increase significantly for most siRNA compared to the siC control, except for siTOPBP1, siMDC1, siUBE2B, siRAD18 and possibly siATR. Of these, only siATR, siUBE2B and siRAD18 show a convincing albeit small increase in DNA:RNA accumulation at RPL13A but this accumulation was not validated at APOE. Overall, the differential between the enrichments obtained + or - RNase H treatment does not seem to be significantly changed for most siRNA compared to the siC control. It could be that the apparent increase in S9.6 signals is primarily due to an increase in the background signal.
• There are significant discrepancies between Figure 1 and Figure 2 that were not discussed: for example, siMDC1 gave one of the strongest increase in the IF signal but gave no increase in the DRIP experiment. On the contrary, siATM lead to an accumulation of DNA:RNA hybrids on Figure 2 but not on Figure 1.
In the end, the accumulation of DNA:RNA hybrids was therefore only rigorously validated at one gene in 3 of the 21 siRNA tested. And the expression of that gene was not tested. Based on the data as they currently stand, it is difficult to conclude that there is a significant genome-wide accumulation of DNA:RNA hybrids in the absence of DDR factors, contrary to what the authors conclude. To support their conclusions, the authors would need to show IF experiments (and their quantifications) after RNase H treatment for at least 3 of the hits identified through the screen (use siTHOC1 and siFANCD2 as positive controls) and they would need to test more than 2 genes by DRIP. It is also possible that DNA:RNA hybrids accumulate in the siRNA conditions tested at other loci than the loci expected based on existing DRIP-seq maps. To test this interesting possibility, would it not be a good idea to characterize the genome-wide pattern of DNA:RNA hybrid accumulation in at least one of their best siRNA condition using DRIP-seq?

2. Figure 3 aims to show that at least some of the DNA damage found in the siRNA identified through the screen is due to the formation/stabilization of DNA:RNA hybrids. However, the data are not very convincing. The authors have performed statistical analysis by comparing to the siC reference for each treatment ({plus minus} RNase H1). However, it seems that the over-expression of RNase H1 might create some comet tail movement on its own. First, the authors should discuss this. The important consequence of this observation is that the over-expression of RNase H1 could 'mask' the comet tail movement that remains after DNA:RNA hybrids have been forcibly disassembled. The authors should therefore also perform statistical analysis on the comparison between + and -RNH1 for each condition. Based on the data that are shown, this difference might only be statistically significant for siUBE2B and perhaps siRad17. This should significantly alter the conclusions drawn by the authors (but not necessarily make them less interesting!).

3. Figure 4 aims to demonstrate that the accumulation of DNA:RNA hybrids upon depletion of DDR components trigger the accumulation of histone marks that reflect the toxicity of DNA:RNA hybrids. There again, some questions arise about the statistical analysis of the data. Although this is not explained in the figure legend, it appears that black stars reflect statistical differences with siC whilst red stars reflect statistical differences between {plus minus} RNH1.
• Fig 4A: does the lack of red star for siATM, siUBE2B and siRAD18 mean that RNH1 over-expression does not significantly reduce the accumulation of H3-S10P positive cells? How do the authors explain that they detected significant hybrid-dependent comet tail movement in all three conditions (Figure 3)?
• Fig 4B: similarly, why the lack of red stars for siATR, siCHK2, siATM, siUBE2B and siRAD18? The differences seem to be significant enough?
• Fig 4B: does it mean that siCHK1 and siRAD17 cells accumulate H3S10P but not H3K9me2 in a hybrid-dependent manner? This should clearly be stated in the text. A double labelling experiment would also clearly validate this interesting observation.
• When the authors say that they counted at least 70 cells, does it mean that they counted at least 70 cells with RNH1 and 70 cells without RNH1 in each experiment (140 cells in total for each condition out of their two experiments)?

4. Figure 5 studies replication fork progression in the different conditions of interest. There again, some questions arise about the statistical analysis of the data.
• Fig 5A: What is the difference between a red and a black star? The authors say that depletion of ATM/CHK2 did not affect fork velocity or track length yet the corresponding data are connected to siC by 3 black or 1 black stars respectively (5A, top) or 3 red stars (EV3, top), suggesting that the differences that the authors observed are significant. What is the genuine conclusion? Regarding the comparison to the '+RNH1' condition, it would be easier to understand the results if the data were

presented on the same graph and the statistical analysis performed for each siRNA to compare + or -RNH1 (see above comment).
• Fig 5B: same comment as for Fig 3B and Fig 5A. The authors should perform statistical analysis on the comparison between + and -RNH1 for each condition.

5. Figure 6: it is not clear whether the nucleolar signal has been removed from the quantifications as was done in Figure 1. Presumably a significant proportion of DNA:RNA hybrids detected in siC should come from transcription and be sensitive to both CORD and DRB. It does not make sense that DRB would be able to reduce hybrid formation in siATR and siUBE2B but not in siC. Please discuss. It might perhaps make more sense to move this figure closer to figure 1 and 2? And to make it a supplementary figure?

6. The authors make the case that R-loops accumulating in the absence of 9-1-1/ATR/CHK1 interfere with the progression of replication forks whilst those accumulating in the absence of ATM/CHK2 do not. To explain these observations, the authors argue that in the absence of ATM/CHK2, R-loops probably form after DNA replication. This should be tested. The authors could synchronize their cells and evaluate the accumulation of DNA:RNA hybrids in the different phases of the cell-cycle by dot-blots or IF. If the authors were able to support their conclusion by experimental evidence, their manuscript would be considerably strengthened.

7. It was reported previously in budding yeast that the formation of R-loops in RNA processing mutants depended on the HR machinery (PMID:23795288). The authors should discuss their data in light of this publication.

Minor points:

1. siTHOC1, which is mentioned as a positive control for DNA:RNA hybrid accumulation by the authors, would not have been selected as a hit in the screen based on the criterion used. Does it mean that the screen was not sensitive enough or that the selection was too stringent? Please discuss.

2. Which of the siFA genes were genuine hits in the screen? Was siFANCD2 one of them?

3. Fig EV2: A RNase H control should be shown for at least one of the siRNA used (siATR-3 for example). Show also representative images. Explain why the A.U on the y-axis are so much greater than on Figure 1B or Figure 6.

Referee #3:

DNA-RNA hybrids represent a major threat to genome integrity by interfering with DNA replication and DSB repair. It is worth noting that despite the importance of DDR pathways in the maintenance of genome integrity, the interplay between genome surveillance pathways and DNA-RNA hybrids has remained largely unexplored. This manuscript by Barroso and colleagues represents a valuable attempt to fill this gap by screening a siRNA library targeting 240 human DDR genes, using the formation of DNA-RNA hybrids as readout. This screen identified genes involved in three different pathways, namely the ATR, ATM and post-replicative repair (PRR) pathways. Interestingly, the authors show that inactivation of these pathways differentially affect the impact of DNA-RNA hybrids on DNA replication, DNA repair and on the accumulation of chromatin compaction marks. From this respect, this study represents an important and timely contribution to the field and should be of general interest to the readers of EMBO Reports. However, this study is very descriptive and lacks mechanistic insights. Moreover, several important issues need to be addressed to strengthen the main conclusions of the manuscript.

Major issues:

1. The first figure (Fig. 1A) is based on the high-throughput imaging of DNA-RNA hybrids using the S9.6 antibody but the way the experiment was performed is somewhat confusing. As I understand, 14 siRNAs were selected for inducing at least a 10% increase in the intensity of S9.6 signal in the nucleus after excluding nucleoli, which are suspected to induce a non-specific signal. However, the siTHOC1 used here as a positive control did not meet these selection criteria. It was

not included either in the validation set (Fig. 1B), so it is not clear how the 14 top-hit candidates compared to validated positive controls. Moreover, the fact that the six additional siRNAs targeting DDC genes included in Fig. 1B (and especially siATR and siCHK2) did not show up in the initial screen (Fig. 1A) is puzzling. As a matter of fact, only 50% of the hits were validated in Fig. 1B (7/14), which is no better than the validation rate of additional DDC genes (3/6), which were not identified in the screen.

2. In data displayed in Fig. 1C, it seems that the intensity of cytoplasmic staining is very strong for some samples (e.g. siRAD1 and siRAD9) compared to others (siC, siTOPBP1) and sometimes overlap with the DAPI staining. The authors need to explain how they ensured that this cytoplasmic signal did not biased automated analyses of nuclear staining intensity. Moreover, it would be important to provide an RNase H control to confirm that the nuclear staining correspond indeed to DNA-RNA hybrids.

3. The fact that DNA breaks detected with the comet assay could be suppressed by RNase H overexpression is interesting as it suggests that indeed, DNA-RNA hybrids are responsible for the accumulation of DNA damage in the absence of key DDR factors. However, the type of comet assay used here is not indicated. Was it performed in alkaline conditions to reveal mostly ssDNA breaks or in neutral conditions to reveal DSBs? This issue needs to be clarified as it has important implications for the interpretation of the data.

4. In general, it is difficult to assess the impact of a given pool of siRNAs on the intensity of the S9.6 signal (microscopy), DRIP (qPCR) and chromatin compaction (IF) because these data are shown independently in different figure panels. Would it be possible to provide an integrated view of this information, with correlation coefficients, to identify siRNAs generating the most consistent responses? The same applies to the effect of RNase H1 on fork progression and stalling (Fig. 5). It would help to show samples +/- RNH1 next to each other and to determine whether the RNase H treatment significantly reduces the effect of siRNAs on forks, even though RNH1-treated samples show no significant difference between each other. This information could be provided as a supplementary figure.

5. The effect of transcription inhibitors on the S9.6 signal in the absence of ATR and UBE2B is interesting (Fig. 6) but the IF images look very different from the examples shown in Fig. 1C, especially regarding the intensity of cytoplasmic signal. What is the reason for this difference?

6. The results regarding the role of PRR factors in preventing the formation of R-loops behind forks is novel and attractive. In essence, it would explain how DNA-RNA hybrids could accumulate at stressed forks without interfering with their progression. The authors would make their case stronger by extracting data on PRR factors from other figures and presenting them together in a set of data that would more directly support the proposed model (Fig. 7).

Minor issues:

1. Ref 40 is incomplete. Is the article still in press?

2. Page 6, line 19: should read "ATR/CHK1 and ATM/CHK2 branches"

3. The legend of figure 6: should read "... signal intensity per nucleus"

1st Revision - authors' response 26 March 2019

Referee #1:

In this study the authors performed a small-scale DDR siRNA screen to identify genes whose knockdown increased R loop levels. They found that reduction in ATR/CHK1, ATM/CHK2, and post-replicative repair proteins caused an increase of DNA-RNA hybrids and DNA damage. H3S10-P and H3K9me2 also accumulated with R loops. Increased fork asymmetry was observed after depletion of proteins in the ATR pathway. The authors proposed that the damage checkpoint and

post-replicative repair are safeguards against R loop induced DNA damage. Although there are some interesting results in this manuscript, many of the experiments are not carefully controlled and interpreted. Several of the main conclusions of this are not yet convincing. This study still needs to be substantially strengthened to be suitable for publication.

Thanks for the constructive suggestions on the manuscript to help make it stronger.

1. In Fig. 1C, depletion of several DDR proteins led to accumulation of cytoplasmic R loops. This is surprising because most if not all of these DDR proteins are exclusively nuclear. Why would loss of DDR proteins cause an increase of cytoplasmic R loops?
Despite the high affinity of the S9.6 antibody for DNA-RNA hybrids, this antibody also recognizes dsRNAs, which are abundant in the cytoplasm due to the mitochondria. We have now performed new S9.6 IF after pre-extraction of the cytoplasm and RNAse III treatment to degrade dsRNA (new Fig 2B and C), confirming the previous results. Furthermore, we have also validated the results by removing the signal after RNAse H treatment and by DRIP, which is a highly specific method of detection of DNA-RNA hybrids, given that putative dsRNA molecules are not amplified by qPCR. Therefore, our conclusion that DDR depletion leads to DNA-RNA hybrid accumulation in the nucleus is solid. Cytoplasmic S9.6 signals are not accounted in this report since they can be a consequence of other dsRNA related phenotypes, which are out of the scope of this study, as we have recently shown (Silva et al, PNAS 2018). New results have been included in Figure 2 and discussed in the text. Thanks.

2. In Fig. 2, the effects of DDR protein knockdown on transcription should be analyzed. If the transcription levels of these genes are altered, the R loops could be affected indirectly. If the transcription of these genes is cell cycle regulated, the effects of siRNAs on the cell cycle should also be tested too.
DNA-RNA hybrid accumulation was tested in several genes. Since the RNA levels were unchanged in the RPL13A in siATR and siUBE2B cells (new results in new Fig EV2A) as now checked by qPCR, we conclude that the differences detected are not due to expression levels.

3. In Fig. 3, the -/+ RNH1 samples for each siRNA should be directly compared (as in Fig. 4). It does not make sense to compare different siRNAs with control siRNA in the presence of RNH1. From the data, it is impossible to tell whether RNH1 worked in control samples.
Performed as requested (new Figure 3).

4. The data in Fig. 5 are confusing. Fork stalling is presumably the reason for fork asymmetry. If R loops increase fork asymmetry, they should affect fork rate too. It is confusing why R loops increased fork asymmetry but did not reduce fork rate.
We certainly agree that fork stalling must be the reason for fork asymmetry. However, fork rate is not necessarily related to fork stalling. We showed this in the past (Salas-Armenteros et al., EMBO J 2017). Both slower and faster forks can stall forks. This is the case of hyperacetylated chromatin, which facilitates the movement of the forks through open chromatin (faster speed) but this does not preclude that more obstacles can can be found on the way, like R loops (higher asymmetry), as we have shown in Salas-Armenteros et al EMBO J (2017). We have added this to the discussion at *page 10*. Thank you.

5. In Fig. 5, the -/+ RNH1 samples for each siRNA should be directly compared. It does not make sense to compare different siRNAs with control siRNA in the presence of RNH1.
Performed as requested (new Figure 5).

6. The data in Fig. 6 only suggest that R loops are formed in a transcription dependent manner, but they don't necessarily suggest that R loops are formed before replication in the absence of DDR proteins.
We agree that the way that it was written might be confusing. These data are now presented as part of new Figure EV2B and C. Thank you.

7. RAD18 has functions other than post-replicative repair. The data on siRAD18 may be over interpreted.
Certainly, RAD18 is also involved in FA. However, we observed no defect in FANCD2 foci formation in siUBE2B cells (new fig EV1A). We have acknowledged the role of RAD18 role in FA

now in the text on *page 6*. Thanks for making this observation.

Referee #2:

Here the authors test the idea that DDR factors could be important for the detection of R-loops and/or to alleviate their toxicity. They carry out a targeted siRNA screen in HeLa cells to identify DDR factors whose depletion leads to the accumulation of DNA:RNA hybrids. DNA:RNA hybrids were detected using the well-characterized S9.6 antibody as well as chromatin marks that were previously shown to be associated with the accumulation of toxic DNA:RNA hybrids. The authors eventually focus their attention on a subset of 9 DDR factors that belong to different DDR pathways and conclude that pre-existing DNA:RNA hybrids could interfere with DNA replication and that DNA:RNA hybrids could also be promoted by unrepaired DSBs. Neither of these two conclusions is particularly novel. Nevertheless, the characterization of the role of DDR factors in preventing or eliminating DNA:RNA hybrids should be of interest to the wider community and could justify publication. However, there are a number of significant issues with the data that need to be addressed first. The demonstration that R-loops accumulate in the absence of some DDR factors (Figures 1,2 and 4) needs to be improved. In addition, the conclusion that depletion of 9-1-1/ATR/CHK1 causes hybrid-dependent fork stalling and DSB needs to be strengthened significantly. Finally, the authors suggest that different DDR mutants might accumulate R-loops in different phases of the cell-cycle (prior or after DNA replication). This would be a very new and very interesting result and the authors have the technical know-how to test this exciting possibility.

Thank you very much for the positive reception of the manuscript and constructive suggestions.

Major points:

1. Figures 1&2 aim to demonstrate that DNA:RNA hybrids accumulate when the expression of a number of DNA Damage Response genes is down-regulated using siRNA. This was done using S9.6 IF to give an overview of the DNA:RNA signal in the cell (Figure 1) and using DRIP to quantify DNA:RNA accumulation at two candidate loci (Figure 2). There are several important issues with these experiments:
• On Figure 1, the specificity of the S9.6 signals was not validated by a treatment with RNase H. In the end, the accumulation of DNA:RNA hybrids was only rigorously tested at two loci, RPL13A and APOE (Figure 2).
As indicated to referee 1, we have included RNAseH controls for the S9.6 IF and actually included a cytoplasm pre-extraction protocol that will get rid of most of the cytoplasmic S9.6 signal and also included a step of RNAseIII treatment that will get rid of all the non-specific dsRNA signal detected by the S9.6 antibody (new Fig 2B and C). We have also included the analysis of two new genes (MIB2 and RHOT2) by DRIP-qPCR (new panels in Fig 2A). Therefore, our conclusion that DDR depletion leads to DNA-RNA hybrid accumulation in the nucleus is solid. Thanks for making us to make conclusions stronger.

• In a number of siRNA conditions, the background signal (as determined by DRIP enrichment after RNase H treatment) seems to increase significantly for most siRNA compared to the siC control, except for siTOPBP1, siMDC1, siUBE2B, siRAD18 and possibly siATR. Of these, only siATR, siUBE2B and siRAD18 show a convincing albeit small increase in DNA:RNA accumulation at RPL13A but this accumulation was not validated at APOE. Overall, the differential between the enrichments obtained + or - RNase H treatment does not seem to be significantly changed for most siRNA compared to the siC control. It could be that the apparent increase in S9.6 signals is primarily due to an increase in the background signal.
Although it is true that the variability of the RNAseH treated samples is very high, likely due to the low amounts that are immunoprecipitated, the increases observed in the all the samples with RNase H respect to the RNase H-treated siC were never significant.

• There are significant discrepancies between Figure 1 and Figure 2 that were not discussed: for example, siMDC1 gave one of the strongest increase in the IF signal but gave no increase in the DRIP experiment. On the contrary, siATM lead to an accumulation of DNA:RNA hybrids on Figure 2 but not on Figure 1.
In the end, the accumulation of DNA:RNA hybrids was therefore only rigorously validated at one

gene in 3 of the 21 siRNA tested. And the expression of that gene was not tested. Based on the data as they currently stand, it is difficult to conclude that there is a significant genome-wide accumulation of DNA:RNA hybrids in the absence of DDR factors, contrary to what the authors conclude. To support their conclusions, the authors would need to show IF experiments (and their quantifications) after RNase H treatment for at least 3 of the hits identified through the screen (use siTHOC1 and siFANCD2 as positive controls) and they would need to test more than 2 genes by DRIP. It is also possible that DNA:RNA hybrids accumulate in the siRNA conditions tested at other loci than the loci expected based on existing DRIP-seq maps. To test this interesting possibility, would it not be a good idea to characterize the genome-wide pattern of DNA:RNA hybrid accumulation in at least one of their best siRNA condition using DRIP-seq?
As stated above, we have now performed S9.6 IF experiments with siATR, siATM and siUBE2B cells after a cytoplasm pre-extraction protocol and RNAseIII treatment that will get rid of all the dsRNA unspecifically detected by the S9.6 antibody (new Fig 2). We have also included the analysis of two new genes (MIB2 and RHOT2) by DRIP-qPCR. Therefore, our conclusion that DDR depletion leads to DNA-RNA hybrid accumulation in the nucleus is solid. We have now addressed the discrepancies in the text page 7. Thanks.

2. Figure 3 aims to show that at least some of the DNA damage found in the siRNA identified through the screen is due to the formation/stabilization of DNA:RNA hybrids. However, the data are not very convincing. The authors have performed statistical analysis by comparing to the siC reference for each treatment ({plus minus} RNase H1). However, it seems that the over-expression of RNase H1 might create some comet tail movement on its own. First, the authors should discuss this.
The important consequence of this observation is that the over-expression of RNase H1 could 'mask' the comet tail movement that remains after DNA:RNA hybrids have been forcibly disassembled. The authors should therefore also perform statistical analysis on the comparison between + and - RNH1 for each condition. Based on the data that are shown, this difference might only be statistically significant for siUBE2B and perhaps siRad17. This should significantly alter the conclusions drawn by the authors (but not necessarily make them less interesting!).
Certainly, and in agreement with previous reports (Salas-Armenteros et al, 2017), overexpression of RNase H causes cells to accumulate more damage (increased tail moments). This is now clearly stated in the text on *page 7*.
We have represented the data accordingly. It can be seen now that only siUBE2B and siRAD18 cells show a tail moment dependent on DNA-RNA hybrids. In addition siRAD9A signal is partially reduced by RNase H overexpression.
We have discussed this in the text (*page 7*) as well as in the new Figure 7A and in the discussion section (*page 11*).
Thanks for asking for this clarification.

3. Figure 4 aims to demonstrate that the accumulation of DNA:RNA hybrids upon depletion of DDR components trigger the accumulation of histone marks that reflect the toxicity of DNA:RNA hybrids. There again, some questions arise about the statistical analysis of the data. Although this is not explained in the figure legend, it appears that black stars reflect statistical differences with siC whilst red stars reflect statistical differences between {plus minus} RNH1.
We apologize for that. As now stated in the Figure legends, black asterisks denote significant increases whereas red stars denote significant decreases. Thank you.

• Fig 4A: does the lack of red star for siATM, siUBE2B and siRAD18 mean that RNH1 over-expression does not significantly reduce the accumulation of H3-S10P positive cells?
How do the authors explain that they detected significant hybrid-dependent comet tail movement in all three conditions (Figure 3)?
• Fig 4B: similarly, why the lack of red stars for siATR, siCHK2, siATM, siUBE2B and siRAD18? The differences seem to be significant enough?
• Fig 4B: does it mean that siCHK1 and siRAD17 cells accumulate H3S10P but not H3K9me2 in a hybrid-dependent manner? This should clearly be stated in the text. A double labelling experiment would also clearly validate this interesting observation.
• When the authors say that they counted at least 70 cells, does it mean that they counted at least 70 cells with RNH1 and 70 cells without RNH1 in each experiment (140 cells in total for each condition out of their two experiments)?
Certainly, we saw a tendency to reduce the accumulation of H3S10-P and H3K9me2 in all samples

but it was only statistically significant in some of them due to the variability between our experiments. We have repeated now the H3K9me2 IF and counted more cells (more than 100 cells with RNH1 in total and more than 200 cells without RNH1, except for siCHK1, in which we counted 158). The number of cells is lower in the experiments with RNase H1 because we count only those cells overexpressing RNase H1 from the whole population.
When considering the whole pathways together, we conclude that DNA damage increases in all DDR depleted cells. However, this increase is independent on hybrids in ATM/CHK2 depleted cells, partially dependent on hybrids in 9-1-1/ATR/CHK1 depleted cells and completely dependent on hybrids in PRR depleted cells. This is now better explained in the text and clarified in the new integrated figure 7A and in the discussion. We believe that the key conclusion is that globally there is a repetitive behaviour of DDR-deficient cells with respect to the accumulation of chromatin compaction signals in association with DNA-RNA hybrids, rather than the individual phenotype conferred by particular silenced genes, thus providing further evidence for the link between harmful R loops and chromatin compaction shown by different labs (Castellano-Pozo et al, EMBO R, García-Pichardo et al, Mol Cell 2017; Colak et al Science 2014; Groh et al, Plos Genetics 2014…). Thanks.

4. Figure 5 studies replication fork progression in the different conditions of interest. There again, some questions arise about the statistical analysis of the data.
• Fig 5A: What is the difference between a red and a black star? The authors say that depletion of ATM/CHK2 did not affect fork velocity or track length yet the corresponding data are connected to siC by 3 black or 1 black stars respectively (5A, top) or 3 red stars (EV3, top), suggesting that the differences that the authors observed are significant. What is the genuine conclusion? Regarding the comparison to the '+RNH1' condition, it would be easier to understand the results if the data were presented on the same graph and the statistical analysis performed for each siRNA to compare + or -RNH1 (see above comment).
• Fig 5B: same comment as for Fig 3B and Fig 5A. The authors should perform statistical analysis on the comparison between + and -RNH1 for each condition.
We apologize for this. As now stated in the Figure legends, black asterisks denote significant increases whereas red stars denote significant decreases. Thank you.
We have now represented the data in the same graph as requested and clarified the text on *page 8*.

5. Figure 6: it is not clear whether the nucleolar signal has been removed from the quantifications as was done in Figure 1. Presumably a significant proportion of DNA:RNA hybrids detected in siC should come from transcription and be sensitive to both CORD and DRB. It does not make sense that DRB would be able to reduce hybrid formation in siATR and siUBE2B but not in siC. Please discuss. It might perhaps make more sense to move this figure closer to figure 1 and 2? And to make it a supplementary figure?
We agree with the referee and this data was moved to supplementary (new Fig EV2B and C) and mentioned earlier in the manuscript text. The nucleolus signal could not be removed in this case since DRB disturbs the morphology of nucleoli as reported (Noaillac-Depeyre et al Biol Cell 1989). The fact that DRB does not reduce the signal in siC cells might be due to this effect of nucleoli distortion. In any case, we think that this is too technical and out of the scope of the paper as it is and we prefer not to discuss this in the text.

6. The authors make the case that R-loops accumulating in the absence of 9-1-1/ATR/CHK1 interfere with the progression of replication forks whilst those accumulating in the absence of ATM/CHK2 do not. To explain these observations, the authors argue that in the absence of ATM/CHK2, R-loops probably form after DNA replication. This should be tested. The authors could synchronize their cells and evaluate the accumulation of DNA:RNA hybrids in the different phases of the cell-cycle by dot-blots or IF. If the authors were able to support their conclusion by experimental evidence, their manuscript would be considerably strengthened.
Certainly, to be able to measure de novo DNA-RNA hybrid formation after depletion, we would need to perform synchronization experiments. However, most of our experiments are done after 72 hours of transfection and it is not possible to keep synchronized cells for such a long period of time. However, to answer to the referee we have measured DNA-RNA hybrids in different cell cycle phases by three different methods (new Figures 6 and EV3B and C). First, we have looked at the correlation between S9.6 IF data and DNA content in pre-extracted and RNase III treated cells. Second, we have measured whole cell S9.6 intensity levels after RNase II treatment and last, we have measured nuclear S9.6 intensity after cell sorting and cytoplasm pre-extraction. In all cases, we

have observed an increase in S9.6 signal in S-G2 versus G1 cells, consistent with our model. This is now described and discussed at the end of the results section (*page 9*).

7. It was reported previously in budding yeast that the formation of R-loops in RNA processing mutants depended on the HR machinery (PMID:23795288). The authors should discuss their data in light of this publication.
This paper reported by the Koshland laboratory argues about the formation of hybrids *in trans*, which if they occur are a minority; most hybrids in cells are formed co-transcriptionally *in cis*. Apart from the fact that determining the co-transcriptional or *in trans* formation of hybrids in DDR-deficient cells is out of the scope of this study, our experiments with cordycepin and DRB indicates that hybrids are formed in a transcription-dependent manner, which implies to be formed *in cis*. In any case, we have changed the text accordingly to avoid confusion.

Minor points:

1. siTHOC1, which is mentioned as a positive control for DNA:RNA hybrid accumulation by the authors, would not have been selected as a hit in the screen based on the criterion used. Does it mean that the screen was not sensitive enough or that the selection was too stringent? Please discuss.
The selection was stringent in order to identify the best hits.

2. Which of the siFA genes were genuine hits in the screen? Was siFANCD2 one of them?
Yes, FANCD2 was one of them, and the other one was FANCA, as stated on page 5.

3. Fig EV2: A RNase H control should be shown for at least one of the siRNA used (siATR-3 for example). Show also representative images. Explain why the A.U on the y-axis are so much greater than on Figure 1B or Figure 6.
RNase H controls were done for the pool of siRNA and is now shown in figure 2B and C. We have also normalized all experiments to the average of the siC and show relative values now in all figures.
Representative images are now shown in new Fig EV2B as requested.

Referee #3:

DNA-RNA hybrids represent a major threat to genome integrity by interfering with DNA replication and DSB repair. It is worth noting that despite the importance of DDR pathways in the maintenance of genome integrity, the interplay between genome surveillance pathways and DNA-RNA hybrids has remained largely unexplored. This manuscript by Barroso and colleagues represents a valuable attempt to fill this gap by screening a siRNA library targeting 240 human DDR genes, using the formation of DNA-RNA hybrids as readout. This screen identified genes involved in three different pathways, namely the ATR, ATM and post-replicative repair (PRR) pathways. Interestingly, the authors show that inactivation of these pathways differentially affect the impact of DNA-RNA hybrids on DNA replication, DNA repair and on the accumulation of chromatin compaction marks. From this respect, this study represents an important and timely contribution to the field and should be of general interest to the readers of EMBO Reports. However, this study is very descriptive and lacks mechanistic insights. Moreover, several important issues need to be addressed to strengthen the main conclusions of the manuscript.

Thank you very much for the positive reception of the manuscript and constructive suggestions.

Major issues:

1. The first figure (Fig. 1A) is based on the high-throughput imaging of DNA-RNA hybrids using the S9.6 antibody but the way the experiment was performed is somewhat confusing. As I understand, 14 siRNAs were selected for inducing at least a 10% increase in the intensity of S9.6 signal in the nucleus after excluding nucleoli, which are suspected to induce a non-specific signal. However, the siTHOC1 used here as a positive control did not meet these selection criteria. It was not included either in the validation set (Fig. 1B), so it is not clear how the 14 top-hit candidates compared to validated positive controls. Moreover, the fact that the six additional siRNAs targeting DDC genes included in Fig. 1B (and especially siATR and siCHK2) did not show up in the initial

screen (Fig. 1A) is puzzling. As a matter of fact, only 50% of the hits were validated in Fig. 1B (7/14), which is no better than the validation rate of additional DDC genes (3/6), which were not identified in the screen.
siTHOC1 was indeed above the control (see blue dot in the graph at figure 1A). However, given the low reliability of the S9.6 quantification in 96-wells, we decided to focus only on the top hit candidates. We have now highlighted in blue all the six DDC additionally selected candidates, and you can see that four of them are above the control in the initial screen, whereas one scored as the control (siATM) and another one just below (siCHK2). This is now better explained in the text (*page 6*). Thank you.

2. In data displayed in Fig. 1C, it seems that the intensity of cytoplasmic staining is very strong for some samples (e.g. siRAD1 and siRAD9) compared to others (siC, siTOPBP1) and sometimes overlap with the DAPI staining. The authors need to explain how they ensured that this cytoplasmic signal did not biased automated analyses of nuclear staining intensity. Moreover, it would be important to provide an RNase H control to confirm that the nuclear staining correspond indeed to DNA-RNA hybrids.
Certainly, the cytoplasmic signal could bias the results but we have included now new S9.6 IF experiments after cytoplasm pre-extraction, getting rid of the cytoplasmic signal and RNAse III treatment to degrade dsRNA (new Fig 2). Furthermore, the results are now validated by their sensitivity to RNAse H treatment, as explained also to the other two referees.

3. The fact that DNA breaks detected with the comet assay could be suppressed by RNase H overexpression is interesting as it suggests that indeed, DNA-RNA hybrids are responsible for the accumulation of DNA damage in the absence of key DDR factors. However, the type of comet assay used here is not indicated. Was it performed in alkaline conditions to reveal mostly ssDNA breaks or in neutral conditions to reveal DSBs? This issue needs to be clarified as it has important implications for the interpretation of the data.
The experiment shown corresponds to alkaline comet assays, which show all kinds of breaks. This is now stated in the text and Methods sections. Thank you.

4. In general, it is difficult to assess the impact of a given pool of siRNAs on the intensity of the S9.6 signal (microscopy), DRIP (qPCR) and chromatin compaction (IF) because these data are shown independently in different figure panels. Would it be possible to provide an integrated view of this information, with correlation coefficients, to identify siRNAs generating the most consistent responses? The same applies to the effect of RNase H1 on fork progression and stalling (Fig. 5). It would help to show samples +/- RNH1 next to each other and to determine whether the RNase H treatment significantly reduces the effect of siRNAs on forks, even though RNH1-treated samples show no significant difference between each other. This information could be provided as a supplementary figure.
The +/- RNH1 data are now represented together as suggested in all figures.
We have also added an integrated view of the DNA-RNA hybrids, DNA damage and fork stalling data to figure 7 as suggested (new panel A). The results are now discussed in the Discussion. Thank you.

5. The effect of transcription inhibitors on the S9.6 signal in the absence of ATR and UBE2B is interesting (Fig. 6) but the IF images look very different from the examples shown in Fig. 1C, especially regarding the intensity of cytoplasmic signal. What is the reason for this difference? That experiment was performed after cytoplasm pre-extraction. This is now stated in the text. Thank you.

6. The results regarding the role of PRR factors in preventing the formation of R-loops behind forks is novel and attractive. In essence, it would explain how DNA-RNA hybrids could accumulate at stressed forks without interfering with their progression. The authors would make their case stronger by extracting data on PRR factors from other figures and presenting them together in a set of data that would more directly support the proposed model (Fig. 7).
As mentioned in point 4, we have also added an integrated view of the DNA-RNA hybrids, DNA damage and fork stalling data to figure 7 as suggested.

Minor issues:

1. Ref 40 is incomplete. Is the article still in press?

Corrected. Thank you.

2. Page 6, line 19: should read "ATR/CHK1 and ATM/CHK2 branches"
Corrected. Thank you.

3. The legend of figure 6: should read "... signal intensity per nucleus"
Corrected. Thank you.

2nd Editorial Decision 19 April 2019

We have now received the 3 enclosed reports on your revised manuscript. As you will see, both referees 1 and 2 point out that the data are not sufficiently strong yet. I therefore think that some more revisions will be required. Normally, I would discuss with the referees what the most important points to be addressed are, but I am leaving for a one week vacation today and wanted to send you the reports now. If you like, we can discuss the revisions in more detail when I will be back in the office, after the 29th of April. This will also give you some time to think about it. I will ask the referees for cross-comments meanwhile.

We will require a complete point-by-point response to all referee concerns with your newly revised manuscript.

REFEREE REPORTS

Referee #1:

The authors have improved the manuscript with new experiments and revision of the text. However, some of the main conclusions are still not adequately supported by the data presented. Some improvements of specific experiments would be helpful to get the paper more convincing.

1. In Fig. 3, RNH1 only suppressed comet tail moment in some knockdown cells. For example, siATM and siCHK2 clearly increased DNA breaks, but the effects were not suppressed by RNH1. The suppression of DNA breaks in siCHK1 cells by RNH1 was modest. The effects of RNH1 in siRAD1/HUS1 and siATR cells were quite different. This is hard to understand. The conclusion that DNA breaks accumulate in the absence of ATM and ATR pathways is not particularly convincing.

2. In Fig. 4A and 4B, the effects of RNH1 were variable in different knockdown cells. There are significant inconsistencies among siRAD1, siRAD9A, and siATR cells, which is hard to explain. There are also inconsistencies between the effects of RNH1 on H3S10-p and H3K9me2 in various knockdown cells.

3. The effects of RNH1 in Fig. 5 were also not quite convincing. The fork asymmetry was suppressed by RNH1 in some knockdown cells but not others (e.g. siRAD17). As siATM siATR, and siRAD18 increased hybrid formation at the same genes (Fig. 2), it is puzzling why their effects on fork asymmetry are so different (Fig. 5B).

4. The interpretations of the DSBs in siATM/siCHK2 cells are somewhat confusing (Fig. 7). On one hand, multiple experiments show that hybrids were increased in these cells at the same loci as in siATR/siCHK1 cells. On the other hand, RNH1 suppression data suggest that the hybrids have completely different effects in siATM and siATR cells. Why? Are the hybrids detected by DRIP in siATR cells irrelevant to DSBs formation and fork asymmetry?

Referee #2:

In this revised manuscript, the authors present a substantial amount of new data. Unfortunately, I believe that even with these new data, the authors still fail to make a compelling case to support their overall conclusions. The demonstration that RNA-DNA hybrids accumulate in the various

DDR mutant conditions studied here remains largely unconvincing because of inconsistencies and conflicting observations (for example siATM would accumulate RNA-DNA hybrids in Fig 2C but not in Fig 1C/6A/6C/EV3). The fact that S9.6 nuclear signals increase significantly upon RNH1 over-expression in the control siRNA (new Figure 2C) is particularly problematic (see below). One is left with the feeling that, as so often with the S9.6 antibody, the authors are trying to make biological sense of differences that are not significant enough to outweigh the experimental noise. There might be a trend in their data but it is not yet very convincing. Nevertheless, I still believe that part of this work is interesting and should be published in EMBO Reports if the authors were to strengthen some of their observations. I would suggest that the best way forward is probably for the authors to focus their experiments on the PRR pathway because their demonstration that PRR mutants accumulate RNase H-sensitive DNA damage is promising and novel.

Major points:

1. Fig 2C: Based on their recent publication in PNAS, the authors say that the nuclear S96 signals are a much better measurement of the amount of RNA-DNA hybrids after cytoplasmic extraction and RNase III digest. Yet the new figure 2C shows that S9.6 signals increase significantly upon RNH1 over-expression and with the same magnitude than in the mutant conditions investigated (siATR/siATM/siUBE2B). This is a major issue that fundamentally questions the validity of the IF approach to quantify RNA-DNA hybrid formation. How do the authors explain this increase? Does it correspond to an unidentified stress product that is unrelated to RNA-DNA hybrids but recognized by the antibody? Or is it that RNH1 over-expression does not remove all RNA-DNA hybrids and even stimulates the formation of RNH1-resistant RNA-DNA hybrids? Why not use exogenous RNase H treatment as reported by the authors previously? If the authors repeat their statistics using the RNH1+ siC as reference, I am concerned that only siATR -RNH1 might show a very small increase in S9.6 signals and it is not clear whether this would be biologically significant. The new results presented in Fig 2C tend to suggest that S9.6 immunofluorescence gives noisy signals and that this experimental noise prevents the drawing of meaningful conclusions. As most of the authors' conclusions rely on immunofluorescence experiments and despite the vast amount of work presented, this major caveat weakens considerably the manuscript.

2. Fig 6. The new data aiming to demonstrate that RNA-DNA hybrids accumulate predominantly in post-replicative cells are not convincing in their current state. If the authors believe that their S9.6 measurements by IF genuinely reflects the abundance of RNA-DNA hybrids in the nucleus, they should show that the weak correlation between S9.6 signal intensity and DNA content is lost upon RNH1 over-expression.
a. Figure 6A/6B: most of the cells that the authors looked at were in G1 and this makes the drawing of a possible correlation between DNA content/S9.6 difficult (because the DNA content of most cells does not change "biologically").
b. If one locus produces RNA-DNA hybrids before replication and continues producing RNA-DNA hybrids after replication, the S9.6 signal should in theory double in G2 compared to G1. So it is not unexpected that S9.6 signals are correlated with the DNA content. To be able to claim that RNA-DNA hybrids accumulate predominantly in G2, would the authors not expect that their G2/G1 ratio should therefore exceed 2?
c. For a given DNA content, one would have expected (based on Fig 1 and Fig 2) that the S9.6 signals should be greater in siATM, siUBE2B and siATR compared to siC. Such a general trend is really not obvious from the plot presented in Figure 6A. Similarly in Figure 6C, the mutant conditions do not appear to accumulate more S9.6 signal than siC, whatever the phase of the cell-cycle.
d. What is the difference between Figure 6 (data from 2C) and Figure EV3? Both analyses were carried out on cytoplasmic extracted, RNase III treated cells according to the text.

3. The addition of the two new genes for the DRIP analysis presented on Figure 2 is not terribly helpful as most of the differences are not statistically significant at those genes. Even if the DRIP assay is more specific than IF, the data do not strongly support the idea of a general increase in RNA-DNA hybrids in the conditions analysed. In addition, a DRIP-negative region should be included like in the previous version of the manuscript to validate the DRIP experiments. Do the DRIP signals in the RNase H-treated samples correspond to a negative region as expected?

Minor points:

• The figures should be clearly labelled to indicate whether the S9.6 IF has been done with or without cytoplasmic extraction and with or without RNase III treatment. In addition, the figures should indicate whether or not the nucleolar signal has been removed for the quantification.

Referee #3:

The authors have addressed all the issues that I raised and in my opinion, the manuscript is now suitable for publication in EMBO Reports.

2nd Revision - authors' response 4 June 2019

The authors have improved the manuscript with new experiments and revision of the text. However, some of the main conclusions are still not adequately supported by the data presented. Some improvements of specific experiments would be helpful to get the paper more convincing.
We acknowledge these constructive comments. We believe that we probably did not make the best job explaining our conclusions, since it is not our intention to conclude that R loops are the unique or main cause of the DNA damage accumulated in DDRdeficient cells. By contrast, we believe that DNA-RNA hybrids accumulate in DDRdepleted cells, which already accumulate DNA damage coming from other sources.
We have taken seriously the comments to try to make things clearer. Thanks.

1. In Fig. 3, RNH1 only suppressed comet tail moment in some knockdown cells. For example, siATM and siCHK2 clearly increased DNA breaks, but the effects were not suppressed by RNH1. The suppression of DNA breaks in siCHK1 cells by RNH1 was modest. The effects of RNH1 in siRAD1/HUS1 and siATR cells were quite different. This is hard to understand. The conclusion that DNA breaks accumulate in the absence of ATM and ATR pathways is not particularly convincing.
We realize that we did not explain this sufficiently. The difference between ATM/CHK2 and 9-1-1/ATR/CHK1 is that whereas ATR responds to stalled forks, ATM responds to DSBs. Therefore, it is not surprising that unresolved DSBs accumulate in the absence ATM/CHK2 regardless of DNA-RNA hybrids. As recently shown, DSBs are preferential sites for the accumulation of DNA-RNA hybrids, because the broken DNA is released of topological constraints allowing hybrids (see Aguilera and Gómez-González NSMB 2017). Consequently, it is expected that DNA-RNA hybrids in the absence of ATM are increased, but these are not suppressed by RNH overexpression. DNA-RNA hybrids in ATM-depleted cells would accumulate preferentially at DSB sites as indeed we propose in the model in Figure 7B. Instead, the modest effect of RNH1 in cells depleted of the 9-1-1/ATR/CHK1 factors indicates that the accumulation of breaks, which are likely a consequence of unresolved stalled forks, is partially dependent on DNA-RNA hybrids that were formed before any fork breakage. We have clarified this better in the text page 7.

This issue is also explained in the discussion (page 11) as follows:
'ATM/CHK2 might also have a role in R loop resolution. However, the accumulation of DNA-RNA hybrids observed in ATM/CHK2-depleted cells does not seem to be a major problem for replication fork progression, as we were not able to detect any increase in fork asymmetry (Fig 6B). Also, the fact that RNase H overexpression had no effect on the number of breaks induced by ATM/CHK2 depletion (Fig 7A) suggests that most breaks occurring in the absence of ATM/CHK2 are independent on DNA-RNA hybrids. Given the views and recent observations supporting that DNA breakage, whether single or double-stranded, is a driving force for DNA-RNA hybrid formation [15,56,63-65], the accumulation of DNA-RNA hybrids in ATM/CHK2 depleted cells might rather be a consequence of such unrepaired DSBs, which would not imply any additional consequences in fork progression.'

2. In Fig. 4A and 4B, the effects of RNH1 were variable in different knockdown cells. There are significant inconsistencies among siRAD1, siRAD9A, and siATR cells, which is hard to explain. There are also inconsistencies between the effects of RNH1 on H3S10-p and H3K9me2 in various knockdown cells.
We do not think these differences reflect any inconsistency. Each siRNA depletes cells differently and RNase H transfections are also not always fully efficient. Furthermore, H3S10-P and H3K9me2 marks are not necessarily related to each other. These signals respond to different molecular events and are controlled by different pathways and describing the specific pathways behind these chromatin modifications is not the objective of this manuscript. Consequently, there is no reason to expect the same behavior in the two of them in quantitative terms, and certainly it would look strange that we would get exactly the same result in all cases. In addition, we have to consider the fact that neither all the components of a protein complex nor all proteins working in a common pathway are expected to have exactly the same phenotypes given that many collateral phenotypes may appear. In other words, we can never be sure of whether some component may have additional roles in parallel processes. The relevant conclusion is the global and consistent result on the accumulation of different chromatin compaction marks in the three DDR branches that we show here to lead to RNA-DNA hybrid accumulation (9-1-1/ATR/CHK1, ATM/CHK2 and PRR).

3. The effects of RNH1 in Fig. 5 were also not quite convincing. The fork asymmetry was suppressed by RNH1 in some knockdown cells but not others (e.g. siRAD17). As siATM siATR, and siRAD18 increased hybrid formation at the same genes (Fig. 2), it is puzzling why their effects on fork asymmetry are so different (Fig. 5B).
Certainly, RNase H overexpression is only able to suppress the fork asymmetry significantly in siRAD9A cells. However, the increased fork asymmetry observed in siRAD1, siRAD9A and siRAD17 is lost after RNase H overexpression.
We have explained this in the text as follows:
'As shown in Fig 5B, we observed a significant increase in fork asymmetry after depletion of RAD1, RAD9A, RAD17 and CHK1. Although not significant, a similar tendency was observed after ATR depletion in agreement with increased fork stalling in the absence of a proper ATR/CHK1 checkpoint response. This increase was lost after RNase H overexpression in siRAD1, siRAD9A and siRAD17-treated cells, supporting that although DNA-RNA hybrids are obstacles to replication fork progression, they are not the only kind of spontaneously occurring obstacles, consistent with our actual knowledge [55].'

Note that in Fig 5A (velocity), all siRAD1, siRAD9A, siRAD17, siCHK1 and siATR show a decrease and these decreases were maintained after RNase H overexpression. However, Fig 5B shows an increase in fork asymmetry in these same factors (siRAD1, siRAD9A, siRAD17 and siATR) that is exclusive for cells that were not overexpressing RNase H.

It is also noteworthy that when we pile up all data from 9-1-1/ATR/CHK1 depleted cells (Fig 7A), the tendency is much more significant.

Regarding the DRIP, it is right that we observed some significant increases in the four genes tested in cells depleted for the three DDR branches. However, these are just reporter genes that are generally used for DRIP analysis and they cannot reflect what would be the DNA-RNA hybrid accumulation profiles genome-wide and whether they are distinct in the different conditions tested. This is beyond the scope of this paper and would require DRIPseq data. We do not try to conclude that these particular DNA-RNA hybrids located in these particular genes are the ones responsible for the fork asymmetry. By contrast, fork asymmetry measures fork stalling in general and not at particular sites. In our view, the fact that we see hybrids but not fork asymmetry in siATM cells responds to the model proposed (Fig 7B). As explained in point 1, in siATM cells, unresolved DSBs would lead to DNA-RNA hybrid formation.

4. The interpretations of the DSBs in siATM/siCHK2 cells are somewhat confusing (Fig. 7). On one hand, multiple experiments show that hybrids were increased in these cells

at the same loci as in siATR/siCHK1 cells. On the other hand, RNH1 suppression data suggest that the hybrids have completely different effects in siATM and siATR cells. Why? Are the hybrids detected by DRIP in siATR cells irrelevant to DSBs formation and fork asymmetry?
Again, this relates to points 1 and 3. Indeed, this is the case. As explained above, the hybrids are formed in different moments. This is the main conclusion of the manuscript that indeed we approach with our model. The relevant issue is not whether hybrids are formed in the same loci but the source of such hybrids. This is the point we try to make in the manuscript and we have tried to make this clearer as explained in point 1.

Referee #2:

In this revised manuscript, the authors present a substantial amount of new data. Unfortunately, I believe that even with these new data, the authors still fail to make a compelling case to support their overall conclusions. The demonstration that RNA-DNA hybrids accumulate in the various DDR mutant conditions studied here remains largely unconvincing because of inconsistencies and conflicting observations (for example siATM would accumulate RNA-DNA hybrids in Fig 2C but not in Fig 1C/6A/6C/EV3). The fact that S9.6 nuclear signals increase significantly upon RNH1 over-expression in the control siRNA (new Figure 2C) is particularly problematic (see below). One is left with the feeling that, as so often with the S9.6 antibody, the authors are trying to make biological sense of differences that are not significant enough to outweigh the experimental noise. There might be a trend in their data but it is not yet very convincing. Nevertheless, I still believe that part of this work is interesting and should be published in EMBO Reports if the authors were to strengthen some of their observations. I would suggest that the best way forward is probably for the authors to focus their experiments on the PRR pathway because their demonstration that PRR mutants accumulate RNase H-sensitive DNA damage is promising and novel.
We appreciate the constructive comments and the clear indication that this is a manuscript to be published in EMBO Reports.

Major points:

1. Fig 2C: Based on their recent publication in PNAS, the authors say that the nuclear S96 signals are a much better measurement of the amount of RNA-DNA hybrids after cytoplasmic extraction and RNase III digest. Yet the new figure 2C shows that S9.6 signals increase significantly upon RNH1 over-expression and with the same magnitude than in the mutant conditions investigated (siATR/siATM/siUBE2B). This is a major issue that fundamentally questions the validity of the IF approach to quantify RNA-DNA hybrid formation. How do the authors explain this increase? Does it correspond to an unidentified stress product that is unrelated to RNA-DNA hybrids but recognized by the antibody? Or is it that RNH1 over-expression does not remove all RNA-DNA hybrids and even stimulates the formation of RNH1-resistant RNA-DNA hybrids? Why not use exogenous RNase H treatment as reported by the authors previously? If the authors repeat their statistics using the RNH1+ siC as reference, I am concerned that only siATR -RNH1 might show a very small increase in S9.6 signals and it is not clear whether this would be biologically significant. The new results presented in Fig 2C tend to suggest that S9.6 immunofluorescence gives noisy signals and that this experimental noise prevents the drawing of meaningful conclusions. As most of the authors' conclusions rely on immunofluorescence experiments and despite the vast amount of work presented, this major caveat weakens considerably the manuscript.
Certainly, RNase H overexpression is toxic and this is not the first time observed. Figure 2C was substituted by a new experiment in which RNase H was used only in vitro and not overexpressed, which might had generated the unexpected increase of signal. The new Figure 2C shows that, as expected, all the observed increases in S9.6 signal are reduced after RNAse H overexpression. Thanks for making us provide cleaner results.

2. Fig 6. The new data aiming to demonstrate that RNA-DNA hybrids accumulate predominantly in post-replicative cells are not convincing in their current state. If the authors believe that their S9.6 measurements by IF genuinely reflects the abundance of RNA-DNA hybrids in the nucleus, they should show that the weak correlation between S9.6 signal intensity and DNA content is lost upon RNH1 over-expression.
We agree that the data only show a weak tendency but this is reproducible after FACS and sorting. Furthermore, we have repeated the experiment in an S/G2 enriched cell (9 hours after thymidine release) and observed the same tendency. Moreover, we were able to see a reduction in these signals by RNase H treatment. This is shown as a figure for referees only (at the end of this document) since we prefer not to add more experiments to the paper that would just reveal the same tendency observed before.

a. Figure 6A/6B: most of the cells that the authors looked at were in G1 and this makes the drawing of a possible correlation between DNA content/S9.6 difficult (because the DNA content of most cells does not change "biologically").
Certainly, most cells in an asynchronous population are in G1 and this correlation might be misleading. We agree with the referee at this point and we have therefore decided to remove panel A from figure 6.

b. If one locus produces RNA-DNA hybrids before replication and continues producing RNA-DNA hybrids after replication, the S9.6 signal should in theory double in G2 compared to G1. So it is not unexpected that S9.6 signals are correlated with the DNA content. To be able to claim that RNA-DNA hybrids accumulate predominantly in G2, would the authors not expect that their G2/G1 ratio should therefore exceed 2?
We were aware of this possibility when submitting the revised version of the manuscript but renounced to include this reasoning to avoid confusion after realizing that the argument is indeed incorrect. We do not expect in G2 to have double amount of hybrids because DNA is not a limiting factor for hybrids. Hybrid formation depends on the RNA produced and it is a very infrequent stochastic event. Once the genome duplicates, transcription is not equally active as in G1. The amount of transcription and RNA produced is not increased after replication, due to gene dosage balance (Voichek et al, Science 2016). Cells do not duplicate the amount of RNA polymerases right away. In addition, in S-G2 cells need to dedicate a lot of effort into histone and rDNA transcription, which takes large part of the transcription and RNA production and machineries. These and likely other transcripts will only affect specific loci and, consequently, hybrids would dilute away in the overall duplicated genome. For this reason, the increase we observe in hybrids in S-G2 is physiologically sound. Furthermore, even though we understand the rational of the point raised by this referee, there is a lack of studies and data to be able to consider the argument that hybrids need to be duplicated in G2. We certainly admit the point and, as said, we were conscious that this possibility could be raised. We now realize that we should have extended the text to convincingly explain the result. We apologize for this. We discussed this in the revised version as follows (page 11):
'Supporting the de novo formation of DNA-RNA hybrids, we detected an increase S9.6 signal in cells in S-G2 (Fig 6). Although it could be argued that gene duplication could double the amount of transcripts, this is known not to be the case due to the gene dosage balance [66]. Furthermore, given that most transcription takes place in G1 and that DNA-RNA hybrid formation is likely a very infrequent stochastic event, genome duplication by itself is not expected to lead to any increase in DNA-RNA hybrids. Consequently, we interpret that the S9.6 enrichment detected is caused by increased formation of DNA-RNA hybrids after replication. Importantly, these DNA-RNA hybrids also lead to genetic instability, but this instability would be replication-independent.'

c. For a given DNA content, one would have expected (based on Fig 1 and Fig 2) that the S9.6 signals should be greater in siATM, siUBE2B and siATR compared to siC. Such a general trend is really not obvious from the plot presented in Figure 6A. Similarly in Figure 6C, the mutant conditions do not appear to accumulate more S9.6 signal than siC, whatever the phase of the cell-cycle.
The data in former Fig. 6A corresponded to Figure 2C and therefore clearly reflected an increase in total S9.6 signals in all siATM, siUBE2B and siATR compared to siC.

Nevertheless, and as stated above, former Figure 6A was now removed.

Figure 6C, as well as 6B and 6D, corresponds to whole cells (not cytoplasm preextracted), measured by FACS (total intensity of S9.6 per whole cell), which is not comparable to IF quantifications (average intensity per nucleus). We pre-treated these cells with RNase III before FACS to minimize unspecific S9.6 signals but we cannot exclude any remaining cytoplasmic signal or remove nucleoli signals as we do in IFs, for example. Nevertheless, Figure 6D shows that the G2/G1 ratio is significantly higher in siATR, siATM and siUBE2B cells than in the control.
For technical reasons, we could not perform this experiment after nuclear preextraction and RNase III treatment, since most cells were lost before the FACS.
Therefore, we repeated the experiment after cytoplasm pre-extraction only and we got the same tendency (Fig EV3B).

d. What is the difference between Figure 6 (data from 2C) and Figure EV3? Both analyses were carried out on cytoplasmic extracted, RNase III treated cells according to the text.
Whereas former Figure 6A came from the data of 2C, Figures 6B, C and D correspond to a different experiment performed with RNase III-treated cells by FACS in a BD FACScalibur cell analyzer and not by IF. Figure EV3B and C corresponds to a different experiment with pre-extracted cells analyzed and sorted in a BD-influx sorter, as indicated in the Methods section. We slightly corrected the text to clarify this point at page 9.

3. The addition of the two new genes for the DRIP analysis presented on Figure 2 is not terribly helpful as most of the differences are not statistically significant at those genes. Even if the DRIP assay is more specific than IF, the data do not strongly support the idea of a general increase in RNA-DNA hybrids in the conditions analysed. In addition, a DRIP-negative region should be included like in the previous version of the manuscript to validate the DRIP experiments. Do the DRIP signals in the RNase Htreated samples correspond to a negative region as expected?
We included the SNRPN results as requested (as a new panel in Figure EV2) and changed the text accordingly at page 6:
'*...previously identified as R loop-prone regions and used as model human genes for these studies [8,25,26,35]. The SNRPN gene was used as a negative control region at which low levels of detection correspond to background (Fig EV2A).*'

Certainly, the DRIP data are not positive in all genes and samples. The important point is that despite looking at the small window of four genes that were just chosen based on previous reports, we can detect a significant DNA-RNA hybrid accumulation in some of our conditions and a tendency in almost all of them. As stated above, understanding genome-wide profiles of DNA-RNA hybrids in all the conditions analyzed is beyond the scope of this paper. In agreement with our model, it is possible that most DNA-RNA hybrids are not accumulated in these particular genes. This result does not contradict the IF data, which are indeed positive (Fig 2C). ATM depletion would likely give rise to DNA-RNA hybrids at DSB sites and not necessarily at these transcribed loci, as proposed in the model.

Minor points:

• The figures should be clearly labelled to indicate whether the S9.6 IF has been done with or without cytoplasmic extraction and with or without RNase III treatment. In addition, the figures should indicate whether or not the nucleolar signal has been removed for the quantification.
We included this as requested. Thanks.

Referee #3:

The authors have addressed all the issues that I raised and in my opinion, the

manuscript is now suitable for publication in EMBO Reports.
Thanks.

Cross-comments from referee 2:

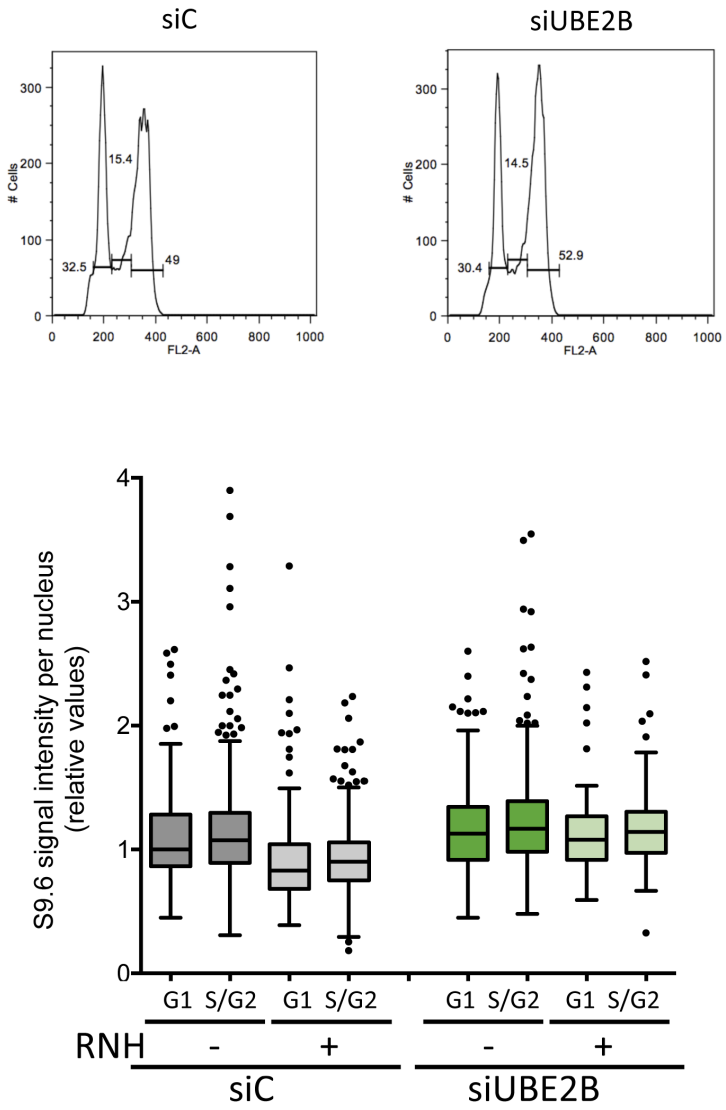
I think that I have been quite exhaustive in my review regarding my concerns about this manuscript but also about its potential qualities/interest. I realise that none of the specific points I have made were raised by the other referees. It would be very interesting for me to see what the other referees think about the particular concerns I have raised, considering that these are recurrent concerns I have about the R-loop literature in general that I think it is time to address. The observation that the S9.6 signals increase after RNH1 over-expression (Fig 2C) does not make sense and questions the validity of most of the observations in this paper.
As explained above, Figure 2C was substituted by a new experiment in which RNase H was used only in vitro and not overexpressed. New Figure 2C shows that all the observed increases in S9.6 signal are reduced after RNase H overexpression.

Cross-comments from referee 3:

I've been through the additional comments raised by Referees #1 and 2. Both of them raise important issues regarding the interpretation of experiments showing variable results after overexpressing RNase H. In a way, this is not surprising as RNase H impacts on various aspects of DNA metabolism, including the processing of Okazaki fragments. It is therefore plausible that this overexpression has differential effects when combined with the depletion of specific DNA repair/checkpoint genes. Overall, there is a trend towards increased S9.6 signal and comet tail moment upon depletion of these factors, but I agree that the rescue by RNase H is more questionable. I guess that repeating all these experiments would not be an option for a second round of revision, so I recommended acceptance, considering that the general trend was convincing enough, despite individual discrepancies. Since this issue is raised by both referees, I would suggest to select a subset of questionable siRNAs and re-test them by other means, including in vitro treatment with bacterial RNase H (as suggested by Referee #2) and inhibition of transcription. Regarding the second point raised by Referee #2, I don't agree that the intensity of S9.6 signal should double in G2 cells but this referee is right to point out that the difference between siC and siATM, siUBE2B and siATR is not very clear. Again, digestion with RNase H in vitro may help here too. Overall, it is a difficult case as the whole study suffers from intrinsic limitations of available tools (S9.6, RNase H...). In my opinion, repeating all the experiments with the same approaches would not necessarily help, but the experiments mentioned above could help.
As stated above, Figure 2C was substituted by a new experiment in which RNase H was used only in vitro and not overexpressed. New Figure 2C shows that all the observed increases in S9.6 signal are reduced after RNAse H overexpression. In vitro RNase H treatment cannot be used to suppress in vivo phenotypes that are a consequence of DNA-RNA hybrids, such as DNA breaks. Therefore, we have used the transcription inhibitor cordycepin to confirm that the partial suppression of the breaks is transcription-dependent (new Fig EV2D).

Regarding the second point, we agree that we cannot expect that the S9.6 signal to be duplicated in G2 cells as we have explained and reasoned extensively to reviewer 2. We also agree that the data only show a weak tendency and we acknowledge the intrinsically difficult technical limitations of available tools. As answered above, this is reproducible after FACS, sorting and also in an S/G2 enriched population (figure for referees only at the end of this document). Note that signals can be reduced by in vitro RNase H treatment in this case as well.

Thanks for these helpful cross-comments.



Cells were blocked in G1/S with 2mM Thymidine for 20 hours, washed three times with PBS, released in fresh medium and harvested after 9 hours to increase the G2 population. Samples were then stained with PI and analysed by flow cytometry to quantify cell cycle populations (top panels). The rest of the population was pre-extracted, treated in vitro with RNAse III or with RNAseIII and RNAse H in vitro and stained with S9.6 antibody. G1 and S/G2 populations were selected by their DNA content (DAPI) and the S9.6 intensity was quantified (bottom panel).

2nd Editorial Decision 25 June 2019

Thank you for the submission of your newly revised manuscript. We have now received the enclosed reports from the referees that were asked to assess it. Both still have a few minor suggestions that I would like you to incorporate before we can proceed with the official acceptance of your manuscript.

A few other changes will also be required:

- please send us a conflict of interest statement

- the panels of figure 3 need to be called out in the manuscript text

- the abstract needs to be written in present tense

- please include scale bars in all microscopy images (some are currently missing)

I have shortened the short summary you sent us, as it was too long (34 words max). Do you agree with this:

DNA-RNA hybrids cause spontaneous DNA damage. DNA damage checkpoint factors, including those of the ATM/CHK2 and ATR/CHK1 pathways, safeguard against DNA-RNA hybrids that spontaneously occur at different stages of the cell cycle.

I look forward to seeing a final version of your manuscript as soon as possible.
Please also indicate where the remaining changes were made in the manuscript file, or include a point by point response with the final version.

REFEREE REPORTS

Referee #1:

The authors have revised the manuscript to explain their model more clearly. Although I agree with the authors that the source of DNA:RNA hybrids may be important for the DSB formation and fork asymmetry in cells compromised for the ATR pathway (not the ATM pathway), it is still quite unclear what exactly is the difference of DNA:RNA hybrids in siATR cells and siATM cells. If the DNA:RNA hybrids in siATM cells come from unsolved DSBs, whereas the hybrids in siATR are the cause of DSBs and fork asymmetry, why are they formed at the same loci? Perhaps something important is still missing in the model. I think that the paper is acceptable for publication -- some of the observations in this paper may be explained by future studies.

Referee #2:

The authors have provided yet more data to support their claims. Fig 2C in particular is now more convincing. I am still concerned that the effects that the authors describe are overall very small, which questions their true biological relevance. However, as pointed out also by the other referees, there might be a trend justifying publication in EMBO Reports, provided that the authors address two very small comments:

1. I am wondering whether there isn't a mistake in the colour of the 3 statistical stars for siC (RNH1 -/+): should they not be red (decrease) rather than black (increase)?
2. I still think that it would be easier for the reader to indicate on the figure panels themselves (rather than in the legend) whether S9.6 IF/FACS have been performed with or without cytoplasmic extraction/in vitro RNase H/III treatment.

2nd Revision - authors' response 28 June 2019

The authors performed all minor editorial changes.

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: Andrés Aguilera
Journal Submitted to: EMBO Reports
Manuscript Number: EMBOR-2018-47250V1

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

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

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

2. Captions

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

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

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ 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?	To estimate sample size when means or medians were calculated, the following formula was used: n= [(Z*S)/E]2, where Z is z-score for 95% of confidence; E is the margin error (5%) and S is Standard Deviation. To estimate sample size when proportions were calculated, the following formula was used: n= [Z2*p*(1-p)]/E2, where p is the expected proportion.
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?	When measuring DNA content of pre-extracted cells; cells with a DNA content higher than 500 arbitrary units were considered aggregates and discarded. In Comet assay, experiments in which the median of the Tail Moment was higher than 30 units were considered outliers and discarded.
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.	NA
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.	Yes. At the microscope, all fields were randomly chosen in DAPI staining. The signals and foci quantification were automatically performed with Metamorph v7.5.1.0 software (Molecular Probes).
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
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Variations between biological replicas are expected to have equal variances and normal distribution. KS normality test was applied to confirm or discard Gaussian distribution in assays with large populations.

USEFUL LINKS FOR COMPLETING THIS FORM

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

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-tun

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 there an estimate of variation within each group of data?	Yes, SD or SEM are represented in every graph, as indicated.
Is the variance similar between the groups that are being statistically compared?	Variations between biological replicas are expected to have equal variances and normal distribution. In IF assays, with data from big populations, variances were different (F-test of variances) but Mann-Whitney tests were applied since the populations did not present Gaussian distribution. When multiple samples were compared, One Way ANOVA analyses of variances were applied.

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).	S9.6 (hybridoma cell line HB-8730), nucleolin (ab50279; Abcam), anti-H3S10P (06-570; Merck), anti-H3K9me2 (07-212; Millipore), Anti-ssDNA from Developmental Studies Hybridoma Bank (DSHB). All the antibodies used in DNA combing assays were described in Bianco JN et al. (2012)
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 source of cell lines was ATCC and they were tested for mycoplasma contamination every six months.

* 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