THE EMBO JOURNAL

Manuscript EMBO-2015-92426

Topoisomerase I-mediated cleavage at unrepaired ribonucleotides generates DNA double-strand breaks

Shar-yin N. Huang, Jessica S. Williams, Mercedes E. Arana, Thomas A. Kunkel and Yves Pommier

Corresponding author: Yves Pommier, National Cancer Institute, NIH

Review timeline:

Submission date:	01 July 2015
Editorial Decision:	17 August 2015
Resubmission:	16 August 2016
Editorial Decision:	22 September 2016
Revision received:	28 October 2016
Accepted:	04 November 2016

Editor: Hartmut Vodermaier

Transaction Report:

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

1st Editorial Decision 17 August 2015

Thank you for submitting your study on Top1-mediated double strand break formation at ribonucleotide incorporation sites. We have now received the reports of three expert reviewers, which you will find copied below. Unfortunately, although the referees appreciate the interest of the topic and the potential importance of your findings, you will see that they all remain unconvinced that especially the presented in vivo evidence is sufficiently direct and definitive to support major conclusions on the physiological mechanism by which Top1 would generate DSBs at unrepaired ribonucleotides. In this light, I am afraid to say we cannot consider the study at present a strong candidate for publication as an EMBO Journal article.

Should future work allow you to obtain such essential more decisive mechanistic insights into in vivo double strand break generation in this context (as well as to address all the other more specific concerns of the reviewers), we would certainly remain open to considering a new version of this study for our journal. However, given that it is currently unclear how such more direct evidence could be obtained and what results these future experiments would yield, I hope you understand that I am not able to invite a formal revision of this manuscript at this stage. In this light, it is only fair to say that such a new version would have to be treated as a new submission with regard to its novelty at the time of resubmission, and only get sent back to our referees if the key concerns would appear to have been conclusively addressed.

Thank you again for having had the opportunity to consider this work for The EMBO Journal. I am sorry that the referee reports do not allow me to come to a more positive conclusion at the present

stage, but hope that you will in any case find our referees' comments and suggestions helpful for your further proceedings with this study.

REFEREE REPORTS

Referee #1:

The manuscript by Huang et al. represents a new addition to the nascent field focused on investigating the significance and consequences of ribonucleotides incorporation in genomic DNA. Previous works have reported that ribonucleotide-containing sites are preferential substrates for the activity of Top1, which will lead to the formation of short deletions in repeated sequences. Through in vitro experiments, Huang et al. report that once Top1 has cleaved the ribonucleotide-containing site, it can frequently cleave the complementary strand generating a DSB. These data are clear and convincing. Then the authors follow up with some experiments to suggest that the DSBs are likely produced also in vivo. This part is less convincing, in the sense that the results are all indirect and there is no direct evidence for the formation of actual DSBs at ribonucleotides. Rad52 foci are not only formed at DSBs but they have been reported also at ssDNA breaks and ssDNA gaps (Lettier et al 2006); gammaH2A derives from activation of Mec1-Tel1 and is not only present after DSB induction.
Overall there are some points that need to be clarified and a few controls that need to be added, but the major problem is that direct evidence of the in vivo formation of DSBs induced by Top1 and dependent upon ribonucleotides should be provided.

Major points
- In figure 1 and S1 the authors analyze the (AT)2 hotspot of the CAN1 gene and scan it by introducing two successive ribonucleotides; they then monitor the cleavage by Top1 within this sequence. I don't understand what is the sense of performing control experiments in conditions that are not the ones that will be used throughout the rest of the manuscript. Why use two ribonucleotides instead of one, as in all the other experiments?
- I don't understand why the chi site is so efficiently cut in Figure 1, even in the absence of a ribonucleotide at that site, and is so inefficiently cut in Figure S1. How reproducible are these assays?
- Figures 1 and S1: from what I understand, Δ4 is the same site as chi, but form Figure S1 the chi site is cleaved by Top1 irrespective of the presence of a ribonucleotide. Why do the authors call Δ4 and chi the same site? This is confusing since the chi site is not dependent upon ribonucleotides.
- Figure 1 and S1. In the legend to figure S1 the authors claim that increasing the number of ribonucleotides abolishes cleavage, on page 3 they present the same data saying that it reduces cleavage only slightly. This again leads to the question of why they used two ribonucleotides instead of one.
- Figure 2. Why are the authors changing the target sequence now? They should use the same sequence for the characterization of the first cleavage and for the rest of the experiments. Similarly they should always use one ribonucleotide, not two as they did in Figure 1
-On page 4, using Top1 trapping drugs, the authors identify only sites Δ1-Δ4 and conclude that ribonucleotides represent a more sensitive way to map Top1 sites. I am not convinced by this interpretation, which may be misleading: how can we exclude that the Δ5 and Δ6 are exclusively targeted when ribonucleotides are present and do not represent Top1 cleavage sites normally? If this is the case it's not a question of sensitivity.
- Figure 2: here the authors should add an internal control showing that Top1 does not induce a single stranded cleavage in the NTS in these conditions (X=dNMP), while it induces a single stranded cleavage in the NTS when X=rNMP. The authors should also show what happens to the NTS strand in the (AT)2 sequence when Top1 cleaves at the chi site in the absence of ribonucleotides. Is the cleavage converted to a DSB or not?
- Discussing figure 2c, the authors refer to a Top1-induced recombination product. I am not sure what the authors mean by recombination here. This is an in vitro reaction carried out by purified Top1. I suppose they refer to ligating the cleaved fragment to the original uncleaved substrate; if so, it should be explained better. This would also mean that Top1 prefers to religate the cleaved 3' end to a different 5' end compared to the one that it just generated and is immediately available. Is there any explanation for this behavior?
- In figure 4a the authors show that cells that accumulate large amounts of ribonucleotides

(pol2M644G rnh201Δ) are slow growing when RAD51 is deleted. This result needs to be interpreted. If, as the authors suggest, Top1 generates DSBs at ribonucleotide-containing positions, the cells should be dead in the absence of Rad51, at least in a rnh201Δ background. They later show that these same mutations are synthetic lethal with rad52Δ. An explanation of this difference should be proposed.
- The pol2-M644G mutation alters the structure of the catalytic site of pol epsilon and this affects the DNA replication process, as seen by FACS analysis. The phenotypes observed when RAD51 is deleted may derive, at least partly, from problems during DNA replication. Please add the wt POL2 controls to figure 4 so that the readers can appreciate this.
- In figure 5 the authors show that mutations that increase ribonucleotide incorporation in the lagging strand, as opposed to the leading strand, are not synthetic lethal with rad52Δ. While I agree that that Rad52 is not essential for viability in these backgrounds, It is clear that the triple mutants are still noticeably sick, which may be due to a partial activity of Top1 on leading strand ribonucleotides. Is this Top1-dependent? Another explanation for the difference between the analysis of the leading strand to that of the lagging strand may be that different polymerase mutants have different ribonucleotide incorporation rates and this may be reflected on the synthetic effects with rad52Δ.
- According to the model, when a ribonucleotide is present, Top1 cleaves at the ribonucleotide-containing site, then two options are possibile: process the intermediate and generate a short deletion, or cleave the opposite strand and generate a DSB. What are the predicted frequencies of the two options?
-How specific for ribonucleotide-containing sites is the DSB induction by Top1? If it is not specific, then we should see high frequency of genomic DSBs when Top1 is active in normal cells that do not accumulate ribonucleotides. If it is specific, what is the mechanism that drives the second cut?

Minor comments
- The Introduction should mention the papers by Ghodgaonkar et al and by Lujan et al showing that ribonucleotides promote faithful MMR; similarly the works of Gunther et al and that of Pizzi et al should be referenced when discussing the effects of the loss of RNase H2 in higher eukaryotes (e.g. page 2 paragraphs 1 and 3)
- It would be easier for the reader if the sequences in figure S1 and that in figure 1 were oriented the same way.
- Vertical double arrow in Figure 1 lane 1 is too long and it covers at least 3 nucleotides instead of 2.
- In figure 1 it is not clear why the efficiency of cleavage at the chi site is variable in different situations. Please elaborate.
- In figure S1 the authors use CPT to lock the cleavable complex and prevent religation, so that the cleavage signal is strong. This suggest that the signal can arise from a combination of increased cleavage and reduced religation. How do we know that longer stretches of ribonucleotides reduce cleavage? They may be facilitating religation. Is this a real possibility?
- The title for figure S1 "Increasing stretches of ribonucleotides abolishes Top1 cleavage sites gradually" should be corrected. I think "increasingly long" or "longer" is more correct that "increasing stretches". "Abolishes" should be "abolish", but I believe that "abolish" is too strong, since cleavage is only abolished in the ribodeoxy sample.
- In figure 2, I would switch panel b and c. First the data, then the model. It makes more sense
- In figure 2c Top1 is missing 1.
- When describing figure 2c the product v should be discussed.
- In figure S2, there are no ribo sites in this experiment, so using the label Δ to indicate Top1-induced nick at ribo-sites is confusing.
- The last character of the legend to figure S3 should be chi.
- When presenting Figure S4, a few words explaining what is the reversal experiment and why it is relevant here are due.
- In the legend to figure S4, when discussing lane 3, the authors talk about site iv; I assume they mean site i. They then discuss site v. This site was also present in figure 2 but was not described. It should be mentioned. Moreover, I don't understand why the relative intensities of the bands is so variable in different gels
- Figure 4 would be more immediately understandable if the fact that all cells were harboring the pol2M644G mutation was indicated
-Figure 5a, lower panel. Microcolonies are barely visible on screen, they will look horrible on paper.

I suggest taking it away or at least moving it to the supplementary figures.
- Figure 5. Please add the green circle rad52Δ legend also to panel c
- Both the synthetic sickness of mutants accumulating elevated levels of ribonucleotides with rad51Δ and the synthetic lethality with rad52Δ were previously reported (Lazzaro et al 2012) and should be referenced
- On page 9 line 17 the authors go back to the "recombination" product. Once again I find this confusing. When talking about in vitro experiments, they seem to refer to the formation of chimeric molecules through religation, when they describe in vivo experiments they seem to refer to a strand-invasion based recombination.
- On page 10 line 12 stimulation is mispelled

Referee #2:

When ribonucleotides incorporated by the replicative DNA polymerases are not removed by the RNase H2 enzyme, they result in genome instability reflected by increased recombination and mutagenesis rates. Several groups have recently shown that these genome instability events are provocated by Top1 cleavage at the rNMP residues and as such, mutagenesis and recombination are Top1-dependent. Recently, the Kunkel group has shown that the Top1 cleavage at rNMPs occurs on the leading strand, reflecting rNTP misincorporation by DNA polymerase epsilon. In this report, the authors present evidence that double strand breaks are mediated by Top1, through two cleavages by Top1, the first at the unremoved rNMP, and the second on the opposite strand.

Others have shown that RNase H2 mutants are hyper-rec and this is dependent on Top1, and the authors here show support for this. Similar to other reports, the authors find increased Rad52 foci, reflecting double strand breaks, in RNase H2 mutants. The authors also show that the increase in Rad52 foci is Top1-dependent, supporting the notion that Top1 is involved in double strand break formation.

The critical issue is the formation of the double strand break after Top1 cleavage at the rNMP residues. The strongest support for this in the paper comes from the in vitro data. All of the in vivo data can be interpreted as Top1 generating the first nick, and the double strand break formation not coming at all from a second cleavage but from replication to the nick. It may be that the first nick occurs during replication, at the fork, when the polymerase complex encounters the rNMP. If not, then one must consider how Top1 is recruited to rNMP residues. This could possibly be through transcription. If it could be shown that double strand breaks are formed in G2 during high transcription, this might provide support for Top1 cleavage in both strands. The authors do cite the relevant studies on transcription and double strand break formation.

It may be the case that most of the second Top1 cleavage is on the same strand as the first cleavage, to remove the cyclic phosphate residue, and that only in some cases is there a cleavage on the opposite strand, as diagrammed in Figure 7. It would be quite interesting to understand whether a cleavage on the opposite strand in vivo is random or is dictated by other features such as secondary structures or chromatin. A related prediction is that the pol2MG mutant should have a higher recombination rate, not just Rad52 foci.

Referee #3:

Huang et al. tested the hypothesis that topoisomerase I (Top1) can cleave at ribonucleotides in DNA and cause a double-strand break (DSB) in the absence of ribonuclease H2, which is the major enzyme cleaving at ribonucleotides incorporated in DNA. Several reports have highlighted the presence of DSBs in the absence of RNase H2, although the mechanism for how such DSBs are generated remains unknown. In this manuscript, biochemical and genetic data are presented to support the hypothesis that Top1 is making DSBs at ribonucleotides unprocessed by RNase H2. Authors also provide evidence for requirement of homologous recombination to repair Top1-induced DNA lesions in the absence of RNase H2 function.

This work has the potential to be a very important contribution to the field. In general, the manuscript is clearly written and experiments are neat and well described, although, the experimental design presents several weaknesses. The presented biochemical and genetic results support the hypothesis that Top1 induces DSBs at ribonucleotides embedded in DNA when RNase H2 is inactive. There is, however, concern that the presented results, while certainly in line with the conclusion taken, are not sufficient to support the conclusion, and additional important control experiments are required. Specific points are discussed below.

Major points
1) In the first chapter of Results it is not clear why the Authors chose substrates containing 2 consecutive ribonucleotides instead of just one ribonucleotide. The Authors do not present any particular logic for this choice. It seems that using a single ribonucleotide would be a more simple experiment, also considering that in successive experiments substrates containing a single ribonucleotide are used. In fact, in Figure 2 a different substrate is used and no explanation is provided why different substrates are needed in experiments of Figures 1 and 2.

2) Results shown in Figure 2c and 2d are not very clear and strong. In Figure 2d, what are the expected sizes of Top1 DSB bands and what are the detected sizes of the Top1 DSB bands? What is the band appearing under the 20 bp size both for - and + Top1? If there is cleavage w/o Top1 on the rU strand, this can interfere with the interpretation of results. The Authors do not show Top1-dependnet cleavage on the NTS strand for this substrate presented in Figure 2. Wouldn't be easier to label the NTS and show that there is cleavage also on the NTS strand? How would bands appear if there is only a nick instead of a DSB? No controls are shown. All this should be addressed.

3) The Authors often use the term <<RER-deficient>> yeast cells for rnh201-null cells. While it is true that rnh201-null cells are RER-deficient (deficient in the Ribonucleotide-Excision Repair mechanism), rnh201-null cells are ALSO deficient in cleavage of R-loops and many other long RNA-DNA hybrids that might be present in the cells. Therefore, the term <<RER-deficient >> as sole phenotype of rnh201-null cells is misleading and not fully correct. Little is done in the study to examine whether DSBs are due to lack of function of RNase H2 at long RNA-DNA hybrids, see below.

4) RNase H2 cleaves at single ribonucleotides in DNA and also at longer stretches of RNA-DNA hybrids in DNA. Differently, RNase H1 cleaves at ribonucleotide stretches of 4 or more in RNA-DNA hybrids. Thus, both RNase H2 and H1 can cleave 4 or long RNA-DNA hybrids such as R-loops. Because R-loops can be a source of DSBs, it is important to test not only the effect of RNase H2 mutations (null RNH201 and rnh201 SFO) but also that of RNase H1 null defect on Rad52 foci formation in order to conclude that Top1 generates DSBs and activates HR at misincorporated ribonucleotides.
The experiment shown in Figure 3 should be done also using the separation of function mutant (SFO)(rnh201-P45D, Y219A) and using rnh1-null mutant, in addition to rnh201 null. Of course it would be interesting to also include the double mutant rnh1 rnh201 null; however single rnh1, as well as SFO, in the opinion of this reviewer, are essential for this study. In Figure 3 experiments, it would be much stronger to see that inactivation of Top1 decreases percentage of cells with Rad52-YFP foci in rnh201-null, as well as in rnh201-SFO but not in rnh1-null cells.

5) Are data presented in Figure 3a normalized for survival frequency? This should be discussed in the Methods or figure legend.

6) Similarly, experiments of Figure 4 should also be done using SFO mutant of RNH201 and using rnh1-null mutant.

7) As well, experiment of Figure 5 should be done using also rnh1-null mutant (crossing rad52 with pol2-M644G rnh1). Because rnh1 rnh201 pol2-M644G is lethal (Lazzaro et al Mol Cell 2012) it is possible that not only pol2-M644G rnh201 rad52 is lethal but also pol2-M644G rnh1 rad52. Using rnh1 would strengthen a lot the results of this experiment of Fig. 5.

8) DSBs could also be caused by Top1 simply nicking (w/o generating a DSB) nearby ribonucleotide sites and then replication fork collapse could lead to a DSB and HR would be required for repair. Which of the presented data exclude this possibility to occur in vivo?

9) The Authors concluded that inactivation of <<RAD51 and RAD52 impairs the ability of cells to cope with Top1-induced damage at ribonucleotides>> and that this is consistent with generation of DSBs by Top1. Wouldn't RAD51 and RAD52 be needed also if Top1 would generate just nicks which would lead to replication fork collapse?

Other points
10) It is stated at the beginning of the first paragraph of Results that << we first confirmed that two consecutive ribonucleotides at the -1 and -2 positions of Top1 binding sites only slightly reduced Top1 cleavage (Fig. S1)>> but no percentage of cleavage are shown. Percentage of cleavage should be shown.

11) In the first chapter of Results the Authors write that they found 5 ribonucleotide-dependent Top1 cleavage in the 16 substrates, indicating that Top1 induces DNA nicks at misincorporated ribonucleotide sites at high frequency (37.5%). Should this be 31.25% (5/16)?

12) Second paragraph of Results << The -1 positions of the four Top1 cleavage sites are underlined in the sequence in Figure 2a (i to iv).>> should be ' The -1 positions of the five Top1 cleavage sites are underlined in the sequence in Figure 2a (i to v)'.

13) Second paragraph of Results: << Top1 induced three bands with the expected size (between 10 and 20 base-pairs)>> can the Authors possibly be more precise and indicate the exact expected sizes?

14) In Figure 1, it would be good to indicate also sizes of bands.

15) In Figure 1, the vertical arrows are often covering part of bands and thus, it would be good if these are moved to the very left border of each gel section to avoid covering bands that are visible in the Top1-less lanes (likely due to alkali sensitivity?).

16) Why in Figure S1 cleavage with Top1 using the Deoxy substrate is much less evident than in Figure 1 at the same strong Top1 site? Is cleavage percentage different? Is the presence of nearby ribonucleotides affecting cleavage at the strong Top1 site? Cleavage percentages should be shown. It would be also helpful to indicate sizes of bands in Fig. S1.

17) What are the Markers in the last lane of Fig. 2c? In this figure, also correct Top to Top1.

18) Last letter in the legend of Fig. S3 should be 'chi symbol' not 'c'.

19) What is the band above Δ1 in Fig. S3?

20) Legend of Fig. S4 <<...Top1 induced cleavage site iv>> should be corrected to 'site i'. Also site iv is not very visible in this figure.

Resubmission - authors' response 16 August 2016

RESPONSE TO REVIEWERS

Comments from the Editor:

"Thank you for submitting your study on Top1-mediated double strand break formation at ribonucleotide incorporation sites. We have now received the reports of three expert reviewers, which you will find copied below. Unfortunately, although the referees appreciate the interest of the topic and the potential importance of your findings, you will see that they all remain unconvinced that especially the presented in vivo evidence is sufficiently direct and definitive to support major conclusions on the physiological mechanism by which Top1 would generate DSBs at unrepaired ribonucleotides. In this light, I am afraid to say we cannot consider the study at present a strong candidate for publication as an EMBO Journal article.

Should future work allow you to obtain such essential more decisive mechanistic insights

into in vivo double strand break generation in this context (as well as to address all the other more specific concerns of the reviewers), we would certainly remain open to considering a new version of this study for our journal. However, given that it is currently unclear how such more direct evidence could be obtained and what results these future experiments would yield, I hope you understand that I am not able to invite a formal revision of this manuscript at this stage. In this light, it is only fair to say that such a new version would have to be treated as a new submission with regard to its novelty at the time of resubmission, and only get sent back to our referees if the key concerns would appear to have been conclusively addressed."

Our answer: Thank you for giving us the opportunity to resubmit and strengthen our manuscript and provide more decisive mechanistic insights into generation of double-strand breaks in vivo. We anticipate the same reviewers will be willing to reevaluate our work. Among the changes and additions, we have developed a new assay based on immune-pull-down experiments for DNA covalently-linked to Top1 followed by S1 nuclease digestion to further provide in vivo evidence for the sequential Top1 cleavage-mediated double strand breaks in cells with unrepaired ribonucleotides. The new experiments have been added in the new Figures 6 and S9. We have also rearranged the manuscript: first we present in vivo evidence for double strand breaks based on yeast genetic experiments (Rad52- and Rad51-deficient strains in RNase H2-, Top1- and Pole-deficient backgrounds), followed by demonstrating for the first time the mechanism for the generation of double-strand breaks by Top1 alone in the presence of a single genomic ribonucleotide using in vitro biochemical assays. Finally, we show that our proposed sequential Top1 cleavage mechanism leads to the formation of double-strand breaks in cells with high levels of unrepaired ribonucleotide using the immune-pull-down experimental scheme mentioned above.

****

Referee #1:

"*The manuscript by Huang et al. represents a new addition to the nascent field focused on investigating the significance and consequences of ribonucleotides incorporation in genomic DNA. Previous works have reported that ribonucleotide-containing sites are preferential substrates for the activity of Top1, which will lead to the formation of short deletions in repeated sequences. Through in vitro experiments, Huang et al. report that once Top1 has cleaved the ribonucleotide-containing site, it can frequently cleave the complementary strand generating a DSB. These data are clear and convincing. Then the authors follow up with some experiments to suggest that the DSBs are likely produced also in vivo. This part is less convincing, in the sense that the results are all indirect and there is no direct evidence for the formation of actual DSBs at ribonucleotides. Rad52 foci are not only formed at DSBs but they have been reported also at ssDNA breaks and ssDNA gaps (Lettier et al 2006); gammaH2A derives from activation of Mec1-Tel1 and is not only present after DSB induction.*
Overall there are some points that need to be clarified and a few controls that need to be added, but the major problem is that direct evidence of the in vivo formation of DSBs induced by Top1 and dependent upon ribonucleotides should be provided."

Our answer: Thank you for judging the biochemical data and model "clear and convincing". We appreciate the detailed comments and constructive criticisms, and your interest in the DSB model. We have clarified the points listed below and added the suggested controls. To provide direct evidence for detecting Top1-induced DSBs opposite to ribonucleotide sites in vivo, we developed a novel assay summarized in the new Figure 6 and S9. Using immuno-pull-down to isolate DNA that is covalently-linked to Top1 followed by S1 nuclease digestion, we now provide evidence for the presence of Top1 at the end of DSBs in yeast strains with elevated ribonucleotides. Together with the fact that Top1 cleavage sites are relatively promiscuous, these data support our model that sequential cleavage by Top1 opposite a newly incorporated ribonucleotide can be a source of DSBs.

"*Major points*
- In figure 1 and S1 the authors analyze the (AT)2 hotspot of the CAN1 gene and scan it by introducing two successive ribonucleotides; they then monitor the cleavage by Top1 within this sequence. I don't understand what is the sense of performing control experiments in conditions that are not the ones that will be used throughout the rest of the manuscript. Why use two ribonucleotides instead of one, as in all the other experiments?"

Our answer: Thank you for raising this point. We have revised the manuscript, added new experiments and reordered the presentation and figures to clarify this point. In the revised manuscript, we show that a single ribonucleotide can induce DSBs in two different sequence contexts [$(AG)_4$ and $(AT)_2$ hotspot] (Fig 5 and S6), consistent with the proposed mechanism (revised Fig 4). To assess the effectiveness and frequency of Top1 ribonuclease activity, we present the ribonucleotide scanning at the $(AT)_2$ hotspot of the *CAN1* gene in Figure S8. We first demonstrate that the ribonuclease activity is not significantly altered with one or two consecutive ribonucleotides (Fig S7). We chose to use two ribonucleotides instead of just one, because it allowed us to increase the coverage of DNA surveyed by two-fold without compromising the sensitivity of the assay. In fact, the cleavage assay with two consecutive ribonucleotides revealed more Top1 sites than three different Top1 trapping poisons combined (Fig S8b, left and right panels). These points have been included in the revised manuscript.

"- I don't understand why the chi site is so efficiently cut in Figure 1, even in the absence of a ribonucleotide at that site, and is so inefficiently cut in Figure S1. How reproducible are these assays?

Our answer: We have adjusted the intensity and contrast of the gel images to comparable levels (Fig S7b and S8b in the revised manuscript) to show that the results are indeed reproducible. The chi site (simply denoted as "site d" in the revised manuscript) is cleaved minimally in a regular deoxyribonucleotide oligonucleotide substrate (Supplemental Fig S7b, lane 2). Yet, it is a preferred site for Top1 in the presence of drugs (Supplemental Fig S8b, right panel) or when a ribonucleotide is present, even at a distance downstream or upstream from this site (Supplemental Fig S8b, left panel). This is likely due to the fact that the interaction of Top1 with "site d" is sensitive to DNA structure containing ribonucleotides. This is noted in the revised manuscript (supplemental figure legend).

"- Figures 1 and S1: from what I understand, Δ4 is the same site as chi, but from Figure S1 the chi site is cleaved by Top1 irrespective of the presence of a ribonucleotide. Why do the authors call Δ4 and chi the same site? This is confusing since the chi site is not dependent upon ribonucleotides."

Our answer: As suggested, we have simplified the annotations of Top1 sites throughout the manuscript. Now we denote the strong Top1 site on $(AT)_2$ TS simply as "site d"; Also see the previous answer for the more in-depth discussion of what factors possibly enhance this particular site.

"- Figure 1 and S1. In the legend to figure S1 the authors claim that increasing the number of ribonucleotides abolishes cleavage, on page 3 they present the same data saying that it reduces cleavage only slightly. This again leads to the question of why they used two ribonucleotides instead of one."

Our answer: As indicated above, we have now revised the manuscript and focused in the main manuscript on data with only one ribonucleotide. It is indeed the case that the ribonuclease activity of Top1 is only effective on short ribonucleotide stretches. The supplemental experiments (Fig S7b, left and middle panels) show that Top1 cleavage is slightly decreased when the number of ribonucleotide in a stretch is increased. Top1 cleavage was completely abolished when the entire construct consisted of only ribonucleotides (Fig S7b, right panel), which is consistent with the previously published results (Sekiguchi & Shuman, 1997 Mol Cell 1: 89-97).

"- Figure 2. Why are the authors changing the target sequence now? They should use the same sequence for the characterization of the first cleavage and for the rest of the experiments. Similarly, they should always use one ribonucleotide, not two as they did in Figure 1"

Our answer: Thank you. We agree with your suggestion, and as indicated above, we have revised the manuscript accordingly and now we expanded our studies to include evidence that a single ribonucleotide lead to Top1-induced DSB in both $(AG)_4$ and $(AT)_2$ target sequences (Fig 5 and S6).

"-On page 4, using Top1 trapping drugs, the authors identify only sites Δ1-Δ4 and conclude that ribonucleotides represent a more sensitive way to map Top1 sites. I am not convinced by this interpretation, which may be misleading: how can we exclude that the Δ5 and Δ6 are exclusively targeted when ribonucleotides are present and do not represent Top1 cleavage sites normally? If

this is the case it's not a question of sensitivity."

Our answer: It is known that drugs trap only a subset of Top1 cleavage sites, requiring specific stacking interactions with the base pairs flanking the Top1 cleavage site, and that different drugs display somewhat different sequence selectivity (see review Pommier and Marchand, 2012 Nat Rev Drug Discov 11: 25-36 and the right panel of Fig S8b). In fact, we do not exclude the possibility that Top1 may bind at site e and site f regardless of the presence of ribonucleotides at these sites (Fig S8). Our point is that none of the Top1 drugs we tested were able to trap Top1 at these 2 sites for detection. Nevertheless, incorporation of ribonucleotides at these positions showed that these sites are targets of Top1 ribonuclease activity (Fig S8). Thus, incorporation of ribonucleotides allowed us to detect Top1 sites that would otherwise not be easily detected. Incorporation of ribonucleotides also showed that the ribonuclease activity of Top1 exhibits broader base sequence preference than the trapping of Top1 by drugs. We have reworded the manuscript and clarified this point in the revised manuscript.

"- *Figure 2: here the authors should add an internal control showing that Top1 does not induce a single stranded cleavage in the NTS in these conditions (X=dNMP), while it induces a single stranded cleavage in the NTS when X=rNMP. The authors should also show what happens to the NTS strand in the (AT)2 sequence when Top1 cleaves at the chi site in the absence of ribonucleotides. Is the cleavage converted to a DSB or not?*"

Our answer: In our previous publication (supplemental Fig S3 of Kim et al, Science, 2013), we showed that Top1 does not induce a nick in the NTS strand of $(AG)_4$ sequence in the absence of ribonucleotide or Top1 poison. As suggested, we have added the control in new experiments showing Top1 cleavage and reversal assay on the $(AG)_4$ sequence when the NTS is radiolabeled on the 3'-end of NTS. When X=dT, there were no CPT-independent sites and CPT-induced Top1 sites were reversible (Supplemental Fig. S5b, lanes 4, 7 and 10). However, when X=rU, the expected product of Top1 ribonuclease activity is both CPT-independent and irreversible (Fig S5b, lanes 12, 15 and 18, site b). We have also carried out cleavage assays for the NTS strand in the $(AT)_2$ sequence. The results are shown in new Supplemental Figure S6, experimental data from both sequencing gel and native gel showed that the strong cleavage site is converted to DSB only when it contains a ribonucleotide at that position.

"- *Discussing figure 2c, the authors refer to a Top1-induced recombination product. I am not sure what the authors mean by recombination here. This is an in vitro reaction carried out by purified Top1. I suppose they refer to ligating the cleaved fragment to the original uncleaved substrate; if so, it should be explained better. This would also mean that Top1 prefers to religate the cleaved 3' end to a different 5' end compared to the one that it just generated and is immediately available. Is there any explanation for this behavior?*"

Our answer: Thank you. We have clarified the manuscript to better explain the molecular mechanism leading to this Top1-induced product migrating slower than the substrate. Figure 5b now includes a cartoon describing the reaction intermediates. Also, in the revised manuscript, we describe the reaction mechanism in Figure 4. As suggested, we have removed the name "recombination product" and renamed it "Intermolecular religation". Top1 can indeed religate to the free 5' end of the DSB that it just generated (Fig 4e to Fig 4d). However, we could not differentiate the product from the original substrate. Furthermore, the concentration of uncleaved substrates and the concentration of the Top1-induced 5'-ends upon DSBs formation are different, making it difficult to compare the efficiency of two different religation events here. This point has been added to the revised manuscript. A reverse arrow has also been added to steps d-e in Figure 4.

"- *In figure 4a the authors show that cells that accumulate large amounts of ribonucleotides (pol2M644G rnh201Δ) are slow growing when RAD51 is deleted. This result needs to be interpreted. If, as the authors suggest, Top1 generates DSBs at ribonucleotide-containing positions, the cells should be dead in the absence of Rad51, at least in a rnh201Δ background. They later show that these same mutations are synthetic lethal with rad52Δ. An explanation of this difference should be proposed.*"

Our answer: Although formation of the Rad51 nucleoprotein filament is a key event in initiation of DNA strand exchange during homologous recombination repair of DSBs, Rad51-independent recombination is also known to occur. This alternate pathway requires Rad52, and deletion of

RAD52 in budding yeast results in a severe defect in all forms of recombination. This may explain why the *pol2-M644G rnh201*D strain is sick in the absence of Rad51 and completely inviable in the absence of Rad52. This is now more carefully described in the text. A similar phenomenon was observed by Lazzaro et al. (2012) in an RNase H-defective strain and is discussed below. They reported that deletion of *RAD51* in *rnh201*D *rnh*1D cells caused increased hydroxyurea-sensitivity, while loss of *RAD52* was lethal in this genetic background.

"- The pol2-M644G mutation alters the structure of the catalytic site of pol epsilon and this affects the DNA replication process, as seen by FACS analysis. The phenotypes observed when RAD51 is deleted may derive, at least partly, from problems during DNA replication. Please add the wt POL2 controls to figure 4 so that the readers can appreciate this."

Our answer: As requested, the FACS profiles for the *POL2-WT* strain are now included as Supplementary Figure S2a. The results demonstrate a mild increase in the percentage of cells in G_2/M in the *rnh201*D *rad51*D strain compared to either single mutant. The fact that deletion of *TOP1* does not significantly impact this altered cell cycle distribution suggests that the threshold of ribonucleotides in the nascent leading strand may not be sufficiently high enough to invoke a quantifiable Top1-dependent effect on cell cycle distribution in the presence of wild type Pol e. We also note that cell cycle distribution of the *pol2-M644G rad51*D strain is not significantly altered compared to the *pol2-M644G* single mutant, suggesting that the phenotypes observed in the *pol2-M644G rnh201*D *rad51*D strain are due to the presence of unrepaired ribonucleotides.

"*- In figure 5 the authors show that mutations that increase ribonucleotide incorporation in the lagging strand, as opposed to the leading strand, are not synthetic lethal with rad52Δ. While I agree that that Rad52 is not essential for viability in these backgrounds, it is clear that the triple mutants are still noticeably sick, which may be due to a partial activity of Top1 on leading strand ribonucleotides. Is this Top1-dependent? Another explanation for the difference between the analysis of the leading strand to that of the lagging strand may be that different polymerase mutants have different ribonucleotide incorporation rates and this may be reflected on the synthetic effects with rad52Δ.*"

Our answer: We agree with the reviewer that the triple mutants (*pol1-L868M rnh201*D *rad52*D and *pol3-L612M rnh201*D*rad52*D) are not healthy, suggesting that DSB repair may also be important in the presence of a high density of unrepaired lagging strand ribonucleotides, although to a lesser extent than in the presence of a high nascent leading strand ribonucleotide load, where *RAD52* is essential. We agree that these lagging strand polymerase mutants have lower ribonucleotide incorporation rates than the *pol2-M644G* mutant and this may contribute to the synthetic effects with inactivating *RAD52*. This is now stated in the discussion. Our preliminary data suggests that the defects of the *pol1-L868M rnh201*D *rad52*D mutant is partially dependent on Top1 and this is an area of research that we are in the process of pursuing.

"*- According to the model, when a ribonucleotide is present, Top1 cleaves at the ribonucleotide-containing site, then two options are possible: process the intermediate and generate a short deletion, or cleave the opposite strand and generate a DSB. What are the predicted frequencies of the two options?*"

Our answer: Interesting question. The generation of a short deletion depends on the Top1-cleavage complex upstream of the ribonucleotide to successfully religate across the DNA gap. Biochemical assays showed that this is a minor product. Genetic experiments also show that this type of short deletions are rare events. On the other hand, Top1 binding on the opposite strand readily generates DSB in the biochemical assays. It is therefore plausible that DSBs occur at much higher frequencies than the short deletions. Further studies are warranted to estimate frequencies of the two pathways. This has now been included in the revised Discussion along with the recent published references from independent groups.

"*-How specific for ribonucleotide-containing sites is the DSB induction by Top1? If it is not specific, then we should see high frequency of genomic DSBs when Top1 is active in normal cells that do not accumulate ribonucleotides. If it is specific, what is the mechanism that drives the second cut?*"

Our answer: As demonstrated in our biochemical assays (Fig 5 and S6), the sequential cleavage on the opposite strand at distances up to 8 nucleotides from the ribonucleotide site readily generates DSBs (see site ii in Fig 5). While we cannot exclude the possibility of Top1 cleavage complexes

formation in the absence of ribonucleotides at these sites, our data clearly demonstrate that the presence of ribonucleotides induces or prolongs the Top1 cleavage complexes at these sites, leading to DSBs. Indeed, without the pre-existing nicks generated by Top1 at the ribonucleotide sites, the Top1 cleavage complexes on the opposite strand do not generate irreversible DSBs.

"*Minor comments*
- The Introduction should mention the papers by Ghodgaonkar et al and by Lujan et al showing that ribonucleotides promote faithful MMR; similarly, the works of Gunther et al and that of Pizzi et al should be referenced when discussing the effects of the loss of RNase H2 in higher eukaryotes (e.g. page 2 paragraphs 1 and 3)"

Our answer: Thank you. These references have been cited where appropriate.

"*- It would be easier for the reader if the sequences in figure S1 and that in figure 1 were oriented the same way.*"

Our answer: As suggested, the sequences have been changed to the same orientation. The ribonucleotide-containing strands are always shown on the bottom throughout the revised manuscript. We also re-oriented and labeled every Top1 site in all sequences in a 5'-to-3' order to be consistent. Thank you.

"*- Vertical double arrow in Figure 1 lane 1 is too long and it covers at least 3 nucleotides instead of 2.*"

Our answer: Thank you. The gel image (Fig S8 in revised manuscript) has been modified.

"*- In figure 1 it is not clear why the efficiency of cleavage at the chi site is variable in different situations. Please elaborate.*"

Our answer: As discussed above, it is likely that the presence of the ribonucleotide at different positions influences the local structure to different degrees, thus impacting Top1 cleavage at "site d" in the revised Figure S8. Indeed, Top1 is known to be sensitive to slight changes in the local structure (reviewed in Pourquier and Pommier 2001 Adv Cancer Res). This point has been added in the figure legend in the revised manuscript.

"*- In figure S1 the authors use CPT to lock the cleavable complex and prevent religation, so that the cleavage signal is strong. This suggests that the signal can arise from a combination of increased cleavage and reduced religation. How do we know that longer stretches of ribonucleotides reduce cleavage? They may be facilitating religation. Is this a real possibility?*"

Our answer: According to our model, if Top1cc is formed at a ribonucleotide site, the back attack of the 2'-hydroxyl can form a nick on DNA and is easily detected. We believe it is unlikely that Top1 forms cleavage on the RNA/DNA hybrid with enhanced religation and yet is not susceptible to the back-attack of the 2'-hydroxyl. Moreover, Sekiguchi and Shuman (1997) also reported that Top1 cannot cleave long stretches of ribonucleotides.

"*- The title for figure S1 "Increasing stretches of ribonucleotides abolishes Top1 cleavage sites gradually" should be corrected. I think "increasingly long" or "longer" is more correct that "increasing stretches". "Abolishes" should be "abolish", but I believe that "abolish" is too strong, since cleavage is only abolished in the ribodeoxy sample.*"

Our answer: We agree, and revised the manuscript accordingly.

"*- In figure 2, I would switch panel b and c. First the data, then the model. It makes more sense*"

Our answer: In the revised manuscript, we first show the yeast data implying DSBs in cells. Then, we present the model as a hypothesis in Figure 4 followed by the experiment validating the hypothesis in Figure 5 and S6. To facilitate the interpretation of the results, we have also included a schematic representation of the different DNA species to the right of the panel in Figure 5b.

"*- In figure 2c Top1 is missing 1.*"

Our answer: Thank you. The figure has been corrected (Fig 5b in the revised manuscript).

"- *When describing figure 2c the product v should be discussed.*"

Our answer: This particular Top1-induced cleavage (product i in Fig 5b-c in the revised manuscript) is not a DSB but a CPT-induced Top1 site that reverses. This is now made clear by including the cleavage and reversal assay as Figure 5c in the revised manuscript.

"- *In figure S2, there are no ribo sites in this experiment, so using the label Δ to indicate Top1-induced nick at ribo-sites is confusing.*"

Our answer: As suggested, we simplified the labeling of different Top1 sites in the revised manuscript.

"- The last character of the legend to figure S3 should be chi."

Our answer: Thank you; we have checked the formatting and fonts throughout the manuscript.

"- When presenting Figure S4, a few words explaining what is the reversal experiment and why it is relevant here are due."

Our answer: Thank you. Covalent Top1 cleavage complexes are readily reversible by adding salt (0.5 M NaCl), whereas DSB do not reverse under these conditions. Therefore, the fact that sites ii, iii, iv and v do not reverse (Fig 5c in the revised manuscript) is consistent with the fact they correspond to DSBs, which is also demonstrated in the native gel (Fig 5d). We have clarified this point in the text accordingly.

"- *In the legend to figure S4, when discussing lane 3, the authors talk about site iv; I assume they mean site i. They then discuss site v. This site was also present in figure 2 but was not described. It should be mentioned. Moreover, I don't understand why the relative intensities of the bands is so variable in different gels*"

Our answer: Thank you for pointing out this editorial mistake. We have now simplified and re-oriented all sequences and the labeling of all Top1 sites. We have also included a brief discussion on this site (site i in the revised manuscript). We have adjusted the contrasts of all gel images in the revised manuscript so that the intensities of different Top1 sites are directly comparable across independent experiments. Note that in the Top1 cleavage and reversal experiment (Fig 5c), the inclusion of CPT, which is not included in Figure 5b, attenuates the intensities of ribonucleotide-induced Top1 sites (Sites iii and iv most notably). However, when one compares the CPT-free lanes, the relative intensities of the bands are consistent (Lane 2 in Fig 5b vs lane 2 in Fig 5c, and Lane 4 in Fig 5b vs lane 11 in Fig 5c).

"- *Figure 4 would be more immediately understandable if the fact that all cells were harboring the pol2M644G mutation was indicated*."

Our answer: The manuscript has been clarified to indicate that all strains harbor the *pol2-M644G* allele in the figure itself in addition to the figure legend (revised Fig 3). In the revised manuscript, we first show the rad52 data in the context of the different DNA polymerase mutations, which are included below each panel in the figure (revised Fig 2).

"-*Figure 5a, lower panel. Microcolonies are barely visible on screen, they will look horrible on paper. I suggest taking it away or at least moving it to the supplementary figures.*"

Our answer: We have modified the image and converted to gray scale. Colonies are now readily visible (Fig 2b). Thank you.

"- *Figure 5. Please add the green circle rad52Δ legend also to panel c*"

Our answer: This legend has now been added to Fig 2c.

"- *Both the synthetic sickness of mutants accumulating elevated levels of ribonucleotides with rad51Δ and the synthetic lethality with rad52Δ were previously reported (Lazzaro et al 2012) and*

should be referenced."

Our answer: Thank you. Lazzaro et al., (2012) demonstrated the synthetic sickness of an *rnh201*Δ *rnh1*Δ *rad51*Δ strain and synthetic lethality of an *rnh201*Δ *rnh1*Δ *rad52*Δ strain (also discussed above). This is now referenced on page 5 of the revised manuscript.

"*On page 9 line 17 the authors go back to the "recombination" product. Once again I find this confusing. When talking about in vitro experiments, they seem to refer to the formation of chimeric molecules through religation, when they describe in vivo experiments they seem to refer to a strand-invasion based recombination.*"

Our answer: As suggested, the terminology has been changed to avoid confusion.

"*- On page 10 line 12 stimulation is misspelled*"

Our answer: Thank you. The manuscript has been carefully edited.

****

Referee #2:

"*When ribonucleotides incorporated by the replicative DNA polymerases are not removed by the RNase H2 enzyme, they result in genome instability reflected by increased recombination and mutagenesis rates. Several groups have recently shown that these genome instability events are provocated by Top1 cleavage at the rNMP residues and as such, mutagenesis and recombination are Top1-dependent. Recently, the Kunkel group has shown that the Top1 cleavage at rNMPs occurs on the leading strand, reflecting rNTP misincorporation by DNA polymerase epsilon. In this report, the authors present evidence that double strand breaks are mediated by Top1, through two cleavages by Top1, the first at the unremoved rNMP, and the second on the opposite strand.*

Others have shown that RNase H2 mutants are hyper-rec and this is dependent on Top1, and the authors here show support for this. Similar to other reports, the authors find increased Rad52 foci, reflecting double strand breaks, in RNase H2 mutants. The authors also show that the increase in Rad52 foci is Top1-dependent, supporting the notion that Top1 is involved in double strand break formation.

The critical issue is the formation of the double strand break after Top1 cleavage at the rNMP residues. The strongest support for this in the paper comes from the in vitro data. All of the in vivo data can be interpreted as Top1 generating the first nick, and the double strand break formation not coming at all from a second cleavage but from replication to the nick. It may be that the first nick occurs during replication, at the fork, when the polymerase complex encounters the rNMP. If not, then one must consider how Top1 is recruited to rNMP residues. This could possibly be through transcription. If it could be shown that double strand breaks are formed in G2 during high transcription, this might provide support for Top1 cleavage in both strands. The authors do cite the relevant studies on transcription and double strand break formation.

It may be the case that most of the second Top1 cleavage is on the same strand as the first cleavage, to remove the cyclic phosphate residue, and that only in some cases is there a cleavage on the opposite strand, as diagrammed in Figure 7. It would be quite interesting to understand whether a cleavage on the opposite strand in vivo is random or is dictated by other features such as secondary structures or chromatin. A related prediction is that the pol2MG mutant should have a higher recombination rate, not just Rad52 foci."

Our answer: We wish to thank the reviewer for acknowledging the novelty of our biochemical data and model. We agree it is possible that double-strand breaks are generated by replication forks as they replicate a strand containing Top1-mediated nicks at ribonucleotides. This possibility has been acknowledged in our revised discussion. Yet, after ribonuclease activity of Top1 cleaves the DNA at a ribonucleotide site, it is also plausible that Top1, which is associated with both transcription and replication complexes, could readily cleave the strand opposite to the existing nick. Top1-induced DSBs were indeed readily observed opposite to single ribonucleotide sites in our biochemical experiments. We present the molecular mechanism in Figure 4 as a hypothesis and we present both

in vitro and in vivo evidence supporting this mechanism. Furthermore, it was recently demonstrated that the *pol2-M644G rnh201*D mutant has an elevated rate of recombination (both loss-of heterozygosity (LOH) and non-allelic homologous recombination (NAHR) (Conover et al. (2015) PMID 26400612). These ribonucleotide-dependent increases in chromosomal rearrangements were strongly dependent on Top1, consistent with a model in which Top1 cleavage events can initiate DSB formation that promotes recombination. This is now discussed in the Introduction. To further establish the possibility that sequential Top1 cleavage events can generate DSBs at ribonucleotide sites, our revision includes results from a novel assay (Fig 6 and S9 in the revised manuscript), providing *in vivo* evidence for Top1-induced DSBs with Top1 linked to the end, via our proposed mechanism (Fig 4). In this novel assay, we pulled down DNA that is covalently linked to Top1 and probed for the position of Top1 on the pulled-down DNA using S1 nuclease digestion. Our data show that covalently-linked Top1 is more frequently found at the end of DSBs in yeast strains lacking RER (Fig 6 and S9), as predicted by our proposed mechanism.

****

Referee #3:

"*Huang et al. tested the hypothesis that topoisomerase I (Top1) can cleave at ribonucleotides in DNA and cause a double-strand break (DSB) in the absence of ribonuclease H2, which is the major enzyme cleaving at ribonucleotides incorporated in DNA. Several reports have highlighted the presence of DSBs in the absence of RNase H2, although the mechanism for how such DSBs are generated remains unknown. In this manuscript, biochemical and genetic data are presented to support the hypothesis that Top1 is making DSBs at ribonucleotides unprocessed by RNase H2. Authors also provide evidence for requirement of homologous recombination to repair Top1-induced DNA lesions in the absence of RNase H2 function.*

This work has the potential to be a very important contribution to the field. In general, the manuscript is clearly written and experiments are neat and well described, although, the experimental design presents several weaknesses. The presented biochemical and genetic results support the hypothesis that Top1 induces DSBs at ribonucleotides embedded in DNA when RNase H2 is inactive. There is, however, concern that the presented results, while certainly in line with the conclusion taken, are not sufficient to support the conclusion, and additional important control experiments are required. Specific points are discussed below."

Our answer: Thank you for judging our study with "the potential to be a very important contribution to the field", and that the "manuscript is clearly written and experiments are neat and well-described". We also wish to thank you for constructive comments, which we have addressed point-by-point below and in the revised manuscript. We believe that the revised manuscript provides further evidence to support that Top1 induces DSBs at unrepaired ribonucleotides. Based on the comments from the editor and the reviewers, we have modified the flow of the manuscript. We begin with *in vivo* evidence for double strand breaks based on yeast genetic experiments (*RAD51*- and *RAD52*-deficient strains). We then propose a biochemical mechanism where sequential Top1 cleavage events at newly incorporated single ribonucleotide sites can lead to double-strand breaks. Next, we show biochemical evidence supporting the proposed mechanism, based on *in vitro* assays in two different sequence contexts. Finally, we show new *in vivo* evidence that the proposed mechanism generates Top1-mediated double strand breaks in cells lacking RNase H2.

"*Major points*
1) In the first chapter of Results it is not clear why the Authors chose substrates containing 2 consecutive ribonucleotides instead of just one ribonucleotide. The Authors do not present any particular logic for this choice. It seems that using a single ribonucleotide would be a more simple experiment, also considering that in successive experiments substrates containing a single ribonucleotide are used. In fact, in Figure 2 a different substrate is used and no explanation is provided why different substrates are needed in experiments of Figures 1 and 2."

Our answer: Thank you. As suggested, we have revised the manuscript by focusing first on the propensity of Top1 to induce DSBs in substrates with only one ribonucleotide in two different sequence contexts (Fig 5 and S6 in the revised manuscript). The 2 consecutive ribonucleotide-scanning experiment (now Fig S8) demonstrate the effectiveness of Top1 ribonuclease activity, even at sites that are not detectable using potent drugs that trap Top1 cleavage complexes (Fig S8).

We used two ribonucleotides instead of just one since it does not significantly impact Top1 binding (Fig S7), yet it allows us to survey 2-fold larger segment of the DNA construct (Fig S8).

"2) *Results shown in Figure 2c and 2d are not very clear and strong. In Figure 2d, what are the expected sizes of Top1 DSB bands and what are the detected sizes of the Top1 DSB bands? What is the band appearing under the 20 bp size both for - and + Top1? If there is cleavage w/o Top1 on the rU strand, this can interfere with the interpretation of results. The Authors do not show Top1-dependent cleavage on the NTS strand for this substrate presented in Figure 2. Wouldn't be easier to label the NTS and show that there is cleavage also on the NTS strand? How would bands appear if there is only a nick instead of a DSB? No controls are shown. All this should be addressed.*"

Our answer: As suggested, the results have been clarified and the sizes of the 3'-labeled oligonucleotides generated by the Top1 DSB have been indicated to the left of panels in the revised Figure 5. The expected band sizes of Top1-induced DSBs are DNA flaps of 11 nt on the NTS strand and 12, 14, 16, or 18 nt on the TS strand, and the observed bands fall between the 10 bp and 20 bp DNA markers as expected. The 20 bp DNA marker was erroneously labeled in the original version of the manuscript. We apologize and have corrected this error. To clarify the interpretation, we have also included a diagram of the expected product size to the right of Figure 5d. The band appearing slightly above the 20 bp size marker is due to the spontaneous hydrolysis of ribonucleotide on the NTS strand, resulting in some of the duplex construct losing a piece of 11 nt. Indeed, if there is only a nick instead of a DSB, the product is expected to migrate similarly to the substrate (30 bp duplex), whereas in the case where the 11 nt-piece is lost, it is expected to migrate similar to this Top1-independent band (DNA flap of 19 nt/30 nt). The band has been labeled with an asterisk and discussed in the figure legend. The interpretation of results is not impacted, as the spontaneous hydrolysis of ribonucleotides gives rise to the same product as Top1 ribonuclease activity at the same site. Finally, we have previously reported the Top1-dependent cleavage on the NTS strand (supplemental Fig S3, Kim et al, Science, 2013). We have also extended the analysis of the NTS strand in a Top1 cleavage and reversal assay, shown in supplemental Figure S5 in the revised manuscript.

"*3) The Authors often use the term <<RER-deficient>> yeast cells for rnh201-null cells. While it is true that rnh201-null cells are RER-deficient (deficient in the Ribonucleotide-Excision Repair mechanism), rnh201-null cells are ALSO deficient in cleavage of R-loops and many other long RNA-DNA hybrids that might be present in the cells. Therefore, the term <<RER-deficient >> as sole phenotype of rnh201-null cells is misleading and not fully correct. Little is done in the study to examine whether DSBs are due to lack of function of RNase H2 at long RNA-DNA hybrids, see below.*"

Our answer: As suggested, we no longer refer to *rnh201*D cells as being RER-deficient. Our manuscript also includes experiments with an RNase H2-Ribonucleotide Excision Defective (*rnh201-RED*) strain, defective in resolving single ribonucleotides yet proficient in resolving stretches of ribonucleotides (illustrated in Fig 1b), which demonstrate that the DSBs observed in the rnh201D and *rnh201-RED* strains are largely attributable to unrepaired single ribonucleotides (see below).

"*4) RNase H2 cleaves at single ribonucleotides in DNA and also at longer stretches of RNA-DNA hybrids in DNA. Differently, RNase H1 cleaves at ribonucleotide stretches of 4 or more in RNA-DNA hybrids. Thus, both RNase H2 and H1 can cleave 4 or long RNA-DNA hybrids such as R-loops. Because R-loops can be a source of DSBs, it is important to test not only the effect of RNase H2 mutations (null RNH201 and rnh201 SFO) but also that of RNase H1 null defect on Rad52 foci formation in order to conclude that Top1 generates DSBs and activates HR at misincorporated ribonucleotides.*
The experiment shown in Figure 3 should be done also using the separation of function mutant (SFO)(rnh201-P45D, Y219A) and using rnh1-null mutant, in addition to rnh201 null. Of course it would be interesting to also include the double mutant rnh1 rnh201 null; however single rnh1, as well as SFO, in the opinion of this reviewer, are essential for this study. In Figure 3 experiments, it would be much stronger to see that inactivation of Top1 decreases percentage of cells with Rad52-YFP foci in rnh201-null, as well as in rnh201-SFO but not in rnh1-null cells."

Our answer: We thank the reviewer for these excellent points, and we have now made considerable effort to more thoroughly address and discuss the ribonucleotide-dependent sources of

DSBs in a cell, including distinction between R-loop- and single ribonucleotide-induced DSBs. We have included the analysis of the *rnh201-SOF* (referred to as *rnh201-RED* in the revised manuscript) in all genetic and phenotypic analyses to examine the DSBs induced by Top1. In the analysis of Rad52-YFP foci and the requirement of Rad52 and the importance of Rad51, we have shown that the source of Top1-induced DNA damage stems from single-unrepaired ribonucleotides using the *rnh201-RED* mutant (see Fig 1,2 and S3). As suggested, we now show that *rnh1*-dificient strain has a small increase in the percentage of cells containing a spontaneous Rad52-YFP focus, but the *rnh201-RED rnh1*Ddouble mutant gave rise to a similar level of spontaneous Rad52-YFP foci as the *rnh201-RED* single mutant, suggesting that single genomic ribonucleotides are the predominant source of Rad52-YFP foci (Fig 1a and 1c). The implications of these results are discussed in depth in the revised manuscript. Thank you.

"*5) Are data presented in Figure 3a normalized for survival frequency? This should be discussed in the Methods or figure legend.*"

Our answer: The data presented are normalized to all cells counted. We have not observed any quantifiable difference in survival frequency of unchallenged cells for any of the strains in this figure. This is now stated in the Methods.

"*6) Similarly, experiments of Figure 4 should also be done using SFO mutant of RNH201 and using rnh1-null mutant.*"

Our answer: In order to maintain focus on DSB formation and repair, the experiments performed in new Figure 3 were all carried out using the *rnh201-RED* allele as well (presented as Fig S3). The same Top1-dependent genome instability phenotypes (increased doubling time, perturbed cell cycle distribution and spontaneous checkpoint activation) observed in the *pol2-M644G rnh201*D *rad51*D cells were observed in the *pol2-M644G rnh201-RED rad51*D strain, in which single ribonucleotides cannot be removed but R-loops can still be processed by RNases H2 and H1. This strongly implies that Top1-cleavage at single ribonucleotides incorporated into DNA by Pol e initiates DNA DSBs that are repaired by HR. This model is further supported by the demonstration that *RAD52* is essential in a *pol2-M644G rnh201-RED* mutant in Figure 2C.

"*7) As well, experiment of Figure 5 should be done using also rnh1-null mutant (crossing rad52 with pol2-M644G rnh1). Because rnh1 rnh201 pol2-M644G is lethal (Lazzaro et al Mol Cell 2012) it is possible that not only pol2-M644G rnh201 rad52 is lethal but also pol2-M644G rnh1 rad52. Using rnh1 would strengthen a lot the results of this experiment of Fig. 5.*"

Our answer: In order to maintain a focus on genome instability caused by single ribonucleotides incorporated by Pol e during replication we feel that the fact that the *pol2-M644G rnh201-RED rad52*D strain is inviable provides strong support for the model that DSBs arising from Top1-cleavage at single unrepaired ribonucleotides in DNA require Rad52-dependent HR for repair. We agree that R-loop mediated replication blocks can also generate DSBs that can be repaired by HR (as reviewed in Aguilera & Garcia-Muse (2012), but feel that analysis of R-loop dependent genome instability is beyond the scope of this study.

"*8) DSBs could also be caused by Top1 simply nicking (w/o generating a DSB) nearby ribonucleotide sites and then replication fork collapse could lead to a DSB and HR would be required for repair. Which of the presented data exclude this possibility to occur in vivo?*"

Our answer: We do not exclude this possibility. Indeed, it is plausible that both possibilities contribute to the formation of DSBs, now presented in the Discussion section. Although the conversion of nicks into DSBs by replication "run-off" or "fork collapse" is not new, the Top1-mediated sequential cleavage opposite to a ribonucleotide nick is presented here for the first time. This model is illustrated in Figure 4, demonstrated biochemically in Figures 5 and S6, and supported by in vivo experiments in Figure 6. It is actually striking that Top1 readily converts a newly incorporated ribonucleotide into DSBs in the two sequences examined (Fig 5 and S6). It is also striking how effectively Top1 ribonuclease activity cleaves DNA at ribonucleotide sites, with an even more relaxed specificity than with drugs that trap Top1 cleavage complexes (Fig S8). Thus, Top1-mediated DSBs at unrepaired single ribonucleotides might be quite frequent. They may also pose the risk of illegitimate religation, as such events were readily detected in our biochemical

assays (Fig 5b and c).

"*9) The Authors concluded that inactivation of <<RAD51 and RAD52 impairs the ability of cells to cope with Top1-induced damage at ribonucleotides>> and that this is consistent with generation of DSBs by Top1. Wouldn't RAD51 and RAD52 be needed also if Top1 would generate just nicks which would lead to replication fork collapse?*"

Our answer: As stated above, we discuss this possibility in the Discussion of our revised manuscript.

"*Other points*
10) It is stated at the beginning of the first paragraph of Results that << we first confirmed that two consecutive ribonucleotides at the -1 and -2 positions of Top1 binding sites only slightly reduced Top1 cleavage (Fig. S1)>> but no percentage of cleavage are shown. Percentage of cleavage should be shown."

Our answer: The percentage of the product band is now listed below the gel image (Revised Fig S7b).

"*11) In the first chapter of Results the Authors write that they found 5 ribonucleotide-dependent Top1 cleavage in the 16 substrates, indicating that Top1 induces DNA nicks at misincorporated ribonucleotide sites at high frequency (37.5%). Should this be 31.25% (5/16)?*"

Our answer: There are 5 ribonucleotide-dependent and 1 ribonucleotide-independent Top1 cleavage sites. The ribonucleotide-independent site is also enhanced by the presence of ribonucleotides, and we calculated the percentage using both types of Top1 cleavage sites.

"*12) Second paragraph of Results << The -1 positions of the four Top1 cleavage sites are underlined in the sequence in Figure 2a (i to iv).>> should be ' The -1 positions of the five Top1 cleavage sites are underlined in the sequence in Figure 2a (i to v)*".

Our answer: The text has been corrected. The sequence has been written in the 5' to 3' orientation for the upper (NTS) strand and the annotations have been clarified in the figure legend.

"*13) Second paragraph of Results: << Top1 induced three bands with the expected size (between 10 and 20 base-pairs)>> can the Authors possibly be more precise and indicate the exact expected sizes?*"

Our answer: The three bands in the native acrylamide gel run slightly above the 10 base pair marker (Fig 5d in the revised manuscript). The expected sizes of Top1-induced DSB bands is a DNA flap of 11 nt on the NTS strand and 12, 14, 16, or 18 nt on the TS strand, consistent with the size of the detected products size. We have also included a diagram of the expected product size to the right of Figure 5d to help the reader.

"14) In Figure 1, it would be good to indicate also sizes of bands."

Our answer: As suggested, we have included the sizes of bands to the side of the gel image (Fig S8b in revised manuscript).

"*15) In Figure 1, the vertical arrows are often covering part of bands and thus, it would be good if these are moved to the very left border of each gel section to avoid covering bands that are visible in the Top1-less lanes (likely due to alkali sensitivity?).*"

Our answer: As suggested, we have edited the figure (now Fig S8b) to avoid covering the bands.

"*16) Why in Figure S1 cleavage with Top1 using the Deoxy substrate is much less evident than in Figure 1 at the same strong Top1 site? Is cleavage percentage different? Is the presence of nearby ribonucleotides affecting cleavage at the strong Top1 site? Cleavage percentages should be shown. It would be also helpful to indicate sizes of bands in Fig. S1.*"

Our answer: The presence of ribonucleotides likely changes the local structure of DNA, thereby affecting Top1 cleavage activity, resulting in the different degree of cleavage at "site d" in different constructs (Fig S8b in revised manuscript). The sizes of the major bands have been added in Supplemental Figure S7 as well.

"*17) What are the Markers in the last lane of Fig. 2c? In this figure, also correct Top to Top1.*"

Our answer: As suggested, we have corrected the typo and included the sizes of the markers to the left side of the figure (Fig 5b in the revised manuscript). Thank you.

"*18) Last letter in the legend of Fig. S3 should be 'chi symbol' not 'c'.*"

Our answer: The manuscript has been carefully edited to correct all formats and fonts. Thank you.

"*19) What is the band above Δ1 in Fig. S3?*"

Our answer: This is likely a non-specific background band (right panel in Fig S8b, above site a). The intensity of the band is weak. Therefore, we did not consider this band as a Top1-cleavage product.

"*20) Legend of Fig. S4 <<...Top1 induced cleavage site iv>> should be corrected to 'site i'. Also site iv is not very visible in this figure.*"

Our answer: We apologize for the confusion and we have simplified the labeling of all Top1 sites in a 5' to 3' orientation in the revised manuscript. We have also adjusted the contrast levels of all gel images to be directly comparable in the revised manuscript. Thank you.

2nd Editorial Decision 22 September 2016

Thank you for submitting a new version of your manuscript on Top1-induced double strand breaks at unrepaired ribonucleotide incorporation sites. It has now been reviewed once more by all three original referees, whose comments are copied below. The experts find the study in general significantly improved, but as you will see, referees 1 and 2 still retain a number of important reservations regarding the experimental data and their interpretation. In this light, should you be able to add all the necessary control experiments requested in particular by referee 1, and to adequately clarify the various other concerns of both reviewers 1 and 2 in your response letter as well as in the manuscript text, then we should be able to consider this work further for ultimate publication in The EMBO Journal.

Thank you again for the opportunity to consider this work for The EMBO Journal. Should you have any questions/comments with regard to this decision, the reports and the revision requirements, please do not hesitate to get back to me.

REFEREE REPORTS

Referee #1:

In this manuscript, Huang et al propose that failure to remove rNMPs from chromosomal DNA leads to the formation of DSBs due to the intervention of Top1. The authors propose that Top1 generates a nick at the rNMP site and then cleaves the opposite strand producing a DSB.

They confirm previous published evidence that yeast cells lacking RNase H2 activity are more likely to generate endogenous Rad52 foci. This increase is lost in the absence of Top1.
Through genetic means, the authors suggest that Rad52 is essential for viability when DSBs are generated by Top1 at rNMPs incorporated on the leading strand.
The conclusion is that the presence of rNMPs in chromosomal DNA leads to the formation of DSBs, which are produced by Top1 through a double incision reaction.
The authors provide in vitro evidence for this reaction and, in this new version of the manuscript, they also add data supporting the formation of DSBs also in vivo.

Overall the authors responded satisfactorily to most of the questions raised in the original reviewing process.
I am still a bit worried by the interpretation of the results obtained through the use of the pol2-M644G mutant, which in my opinion has a clear defect in replication. Nonetheless, considering the whole body of evidence presented in this manuscript I think that the model proposed is sufficiently

supported.
My major concerns regarding the in vivo experiment that has been added to this new version are: a) the whole point of the experiment is to show that Top1 is covalently bound to DSB ends of DNA when rNMPs are not removed form chromosomes. The experiment shown lacks the critical control proving that the recovered DNA fragments are indeed bound to Top1. A non tagged Top1 should have been used as control to prove this point. b) the fact that Top1 is bound to DSB ends as opposed to nicks does not really prove the model. It does not tell us that Top1 first nicks one strand than cleaves the second one. The DSBs could arise also if Top1 generated a nick and the complementary strand broke opposite the nick (or if the rNMP opposite the nick was hydrolysed).

Some concern in the interpretation of the data also arise form the fact that some experiments, repeats of experiments presented by the Kunkel group in Williams et al Mol Cell 2013, give different results:
Fig2 in Williams et al shows that in rnh201Δ cells RNR3 is phosphorylated, while this is not so in Fig3c in this manuscript
Fig3 in Williams et al shows that top1Δ partially rescues the cell cycle defects observed in pol2-M644G rnh201Δ cells, while in fig.S2b in this manuscript this is not the case.

Figure 2c is not convincing as shown. Clearly the plate was incubated for a short time, in fact the colonies are quite small, especially when compared to those in panel a. If the pol2-M644G rnh201-RED rad52Δ were sick (as opposed to dead) as those in dissecting 4-12 in panel a, with the short incubation time colonies would not be visible anyway. So, in this experiment, there is no way to distinguish dead from sick. Longer incubations times should be presented.

In Figure 5d it is quite obvious that the rU samples are overloaded compared to the dT samples, so it is hard to conclude that dT samples do not generate DSB bands.

I think that the mentioned controls should be fairly easy to get and hopefully they will just make the paper strong and will not give unexpected results.

Referee #2:

In this revised manuscript the authors have performed additional experiments and provided more evidence for DSB formation by Topoisomerase I at unrepaired ribonucleotides. They have also answered many of the detailed questions brought up by the reviewers. The manuscript is much improved. The use of the rnh201-RED allele helps in pointing to single ribonucleotides in vivo as being the initiating factor for DSB formation.

However, while the authors correctly acknowledge that some strain combinations are viable but are far from normal, they do not address this. I am not asking for additional experiments. The discussion could try to address this problem. Clearly the in vivo situation is more complex and multiple processing of ribonucleotides on both the leading and lagging strands can lead to genomic instability. While the Conover paper reported that pol2-M644G increased genome instability with their particular assay and was reduced by a top1 mutation, the accompanying paper by O'Connell et al did not find such an increase. This paper should be cited and discussed. Another paper by Epshtein et al (Microbial Cell 3, 248, 2016) also did not find an increase in genome instability recombination events in RNase H2-deficient pol2-M644G cells. This is to point out that there are multiple ways to process ribonucleotides in DNA that lead to recombination-initiating events, and indeed not all may be DSBs, but may be Top1-dependent.

Figure 2a shows that top1 mutation rescues viability of pol2-M644G rnh201 cells, but viability is far from normal. This needs some discussion. Does this mean that there are still breaks but above the threshold for no cells dividing? Or are breaks repaired but in a manner that gives rise to rearranged genomes? Or is this related to the general malaise of the pol2-M644G strain? Similarly, the pol1 and pol3 mutations in the rnh201 strain are not normal. This could suggest that processed or tolerated ribonucleotides are genome-destabilizing in some manner.

I would also like to see an expanded discussion of the model in Figure 4. If the ribonucleotide is in a repeat sequence, then as has been shown in vivo and in vitro, a second Top1 cleavage near the cyclic

phosphate and slippage can result in a deletion without a DSB being formed. How the second cleavage and realignment are coordinated is not entirely understood. However, if the ribonucleotide is not in a repeat, which may be more frequent, there may be no slippage realignment so a cleavage could occur on the same strand as the cyclic phosphate or on the opposite strand to give a DSB. Religation and processing (the cyclic phosphate removal in the left of the f panel is not addressed) could give some type of gap filling or recombination. Based on the ribonucleotide misincorporation rate and the reported recombination rates (which do not detect all recombination events), it may be that relegation/gap filling is more prevalent.

Minor note:
Some of the data are redundant, such as tables and graphs of the same data set, especially in the supplementary material (eg graphs and tables).

Referee #3:

Overall, the manuscript has improved a lot. Results are much stronger.
Beautiful work.

2nd Revision - authors' response 28 October 2016

Revision reviewer comments

Referee 1:

In this manuscript, Huang et al propose that failure to remove rNMPs from chromosomal DNA leads to the formation of DSBs due to the intervention of Top1. The authors propose that Top1 generates a nick at the rNMP site and then cleaves the opposite strand producing a DSB.

They confirm previous published evidence that yeast cells lacking RNase H2 activity are more likely to generate endogenous Rad52 foci. This increase is lost in the absence of Top1.
Through genetic means, the authors suggest that Rad52 is essential for viability when DSBs are generated by Top1 at rNMPs incorporated on the leading strand.
The conclusion is that the presence of rNMPs in chromosomal DNA leads to the formation of DSBs, which are produced by Top1 through a double incision reaction.
The authors provide in vitro evidence for this reaction and, in this new version of the manuscript, they also add data supporting the formation of DSBs also in vivo.

Overall the authors responded satisfactorily to most of the questions raised in the original reviewing process.
I am still a bit worried by the interpretation of the results obtained through the use of the pol2-M644G mutant, which in my opinion has a clear defect in replication. Nonetheless, considering the whole body of evidence presented in this manuscript I think that the model proposed is sufficiently supported.
My major concerns regarding the in vivo experiment that has been added to this new version are: a) the whole point of the experiment is to show that Top1 is covalently bound to DSB ends of DNA when rNMPs are not removed form chromosomes. The experiment shown lacks the critical control proving that the recovered DNA fragments are indeed bound to Top1. A non tagged Top1 should have been used as control to prove this point. b) the fact that Top1 is bound to DSB ends as opposed to nicks does not really prove the model. It does not tell us that Top1 first nicks one strand than cleaves the second one. The DSBs could arise also if Top1 generated a nick and the complementary strand broke opposite the nick (or if the rNMP opposite the nick was hydrolysed).

Response: We thank the reviewer for the constructive and positive comments, for judging our revision as having responded satisfactorily to the questions raised in the original review process, and for finding our model sufficiently supported. To answer the concerns regarding the *in vivo* pull-down experiments, we have further included a control experiment, now in Figure EV4, panel B, where we further treated the pulled-down DNA with TDP1. As described in Figure 6A-B, Top1

cleavage complexes leave a small peptide covalently attached to the DNA after proteinase K digestion of the pulled-down DNA. We reasoned that the small peptide should be efficiently removed by TDP1. As the gel image in Figure EV4B shows, TDP1 digestion further increased the gel mobility of the pulled-down DNA, confirming the DNA is covalently attached to Top1. TDP1 treatment had no effect on the control DNA from the same sample, purified after the Top1 covalently-attached DNA has been removed by the pull-down procedure.
In response to the second point ("b") regarding additional mechanisms by which Top1 could give rise to DSBs, the reviewer is correct that additional mechanisms can generate DSBs at ribonucleotide sites in the genome. However, the DSBs would not be covalently attached to Top1 if they are generated by Top1 first nicking at rNMP and the complementary strand subsequently breaks through other means. Similarly, the DSB generated in the case the rNMP opposite a nick is spontaneously hydrolyzed would not bear a Top1 at the end. Our analysis in Figure 6 showed that the proportion of the Top1 covalently attached to DNA DSB ends increases in rnh201-null strain, supporting the proposed mechanism shown in Figure 4. The possibility that Top1 could generate DSBs by additional mechanisms in the presence of ribonucleotides has been included in the Discussion of the current revised manuscript.

Some concern in the interpretation of the data also arise from the fact that some experiments, repeats of experiments presented by the Kunkel group in Williams et al Mol Cell 2013, give different results:
Fig2 in Williams et al shows that in rnh201Δ cells RNR3 is phosphorylated, while this is not so in Fig3c in this manuscript.

Response: We apologize for the confusion. Figure 2 in Williams et al., 2013, shows increased Rnr3 protein level in a *pol2-M644G rnh201*D strain that is Top1-dependent. The immunoblotting experiment in Figure 3C of this manuscript was performed using a different (and less-sensitive) lot of commercially-available antibody against Rnr3 and shows only slight elevation of Rnr3 protein expression (lane 2) that is Top1-dependent (as the band is less-intense in lane 5). We were unsatisfied with the reduced sensitivity of this lot of Rnr3 antibody and therefore had an antibody raised against Hug1, a protein that is present at high levels during checkpoint activation in yeast. The immunoblotting experiment in Figure 3C clearly shows elevated Hug1 protein level in the *pol2-M644G rnh201*D strain, and Hug1 protein is almost undetectable in *pol2-M644G rnh201*D *top1*D cells. A similar Top-dependent increase in Hug1 protein level is also demonstrated for the *pol2-M644G rnh201-RED* strain in Figure EV1, suggesting that Top1-dependent checkpoint activation is triggered by processing of unrepaired single ribonucleotides.

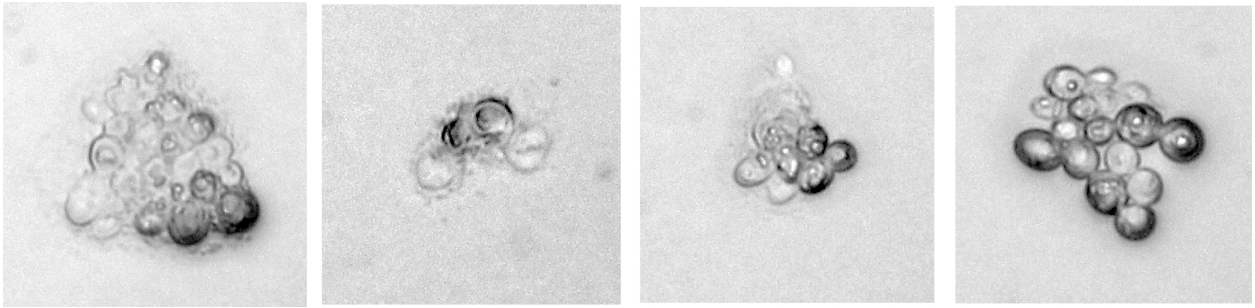
Fig3 in Williams et al shows that top1Δ partially rescues the cell cycle defects observed in pol2-M644G rnh201Δ cells, while in fig.S2b in this manuscript this is not the case.

Response: We thank the reviewer for careful analysis of the present and published data, and again apologize for the confusion. The histograms displayed in Appendix Fig S2B are representative images from two independent experiments. Although it is difficult to see in the histograms, the quantitative data from two independent experiments (plotted as mean values +/- standard error) that corresponds to Appendix Fig S2B is presented in Figure 3B. The legend to Figure 3B now refers the reader to Appendix Fig S2B for representative histograms. The quantitation in Figure 3B demonstrates that Top1-deletion partially rescues the cell cycle distribution defect of *pol2-M644G rnh201*D cells. There is a modest but statistically significant reduction in the percentage of cells in G_2/M in the *pol2-M644G rnh201*D *top1*D strain compared to the *pol2-M644G rnh201*D mutant ($p = 0.01$; unpaired Students *t*-test).

Figure 2c is not convincing as shown. Clearly the plate was incubated for a short time, in fact the colonies are quite small, especially when compared to those in panel a. If the pol2-M644G rnh201-RED rad52Δ were sick (as opposed to dead) as those in dissecting 4-12 in panel a, with the short incubation time colonies would not be visible anyway. So, in this experiment, there is no way to distinguish dead from sick. Longer incubations times should be presented.

Response: We thank the reviewer for raising this point. The plates in the images in Figure 2C were incubated for 4 days at 30˚C following dissection before being photographed, while those in panel A were incubated for 5 days (although the plates in panel A were also photographed at 4 days and the small colonies were already visible at this point). We apologize for not having images of the plates

from 5-day incubation in panel C, but we observed no visible spore colonies even after extended growth of more than 6 days, at which point we genotyped the strains. Microscopic analyses indicated that the spores germinated but did not divide more than a few times. As was seen for the *pol2-M644G rnh201*D *rad52*D haploid strain in Figure 2B, the *pol2-M644G rnh201-RED rad52*D cells arrested as large G_2/M cells. We now make note of this in the legend to Figure 2C. Representative images are shown here:



In Figure 5d it is quite obvious that the rU samples are overloaded compared to the dT samples, so it is hard to conclude that dT samples do not generate DSB bands.

Response: Experiments have been repeated and we have not observed DSB bands in the dT sample above background levels. Further exposure of the gel image in Figure 5D also does not show bands in the dT sample above background level. As suggested, the gel image in Figure 5D has been replaced to clarify this point.

I think that the mentioned controls should be fairly easy to get and hopefully they will just make the paper strong and will not give unexpected results.

Response: We thank the reviewer again for the constructive and insightful comments. The additional controls included in the current revision strengthen the manuscript. Thank you.

Referee 2:

In this revised manuscript the authors have performed additional experiments and provided more evidence for DSB formation by Topoisomerase I at unrepaired ribonucleotides. They have also answered many of the detailed questions brought up by the reviewers. The manuscript is much improved. The use of the rnh201-RED allele helps in pointing to single ribonucleotides in vivo as being the initiating factor for DSB formation.

Response: We thank the reviewer for recognizing we have answered many detailed questions brought up by the reviewers in the initial review, and for finding our manuscript much improved.

However, while the authors correctly acknowledge that some strain combinations are viable but are far from normal, they do not address this. I am not asking for additional experiments. The discussion could try to address this problem. Clearly the in vivo situation is more complex and multiple processing of ribonucleotides on both the leading and lagging strands can lead to genomic instability. While the Conover paper reported that pol2-M644G increased genome instability with their particular assay and was reduced by a top1 mutation, the accompanying paper by O'Connell et al did not find such an increase. This paper should be cited and discussed. Another paper by Epshtein et al (Microbial Cell 3, 248, 2016) also did not find an increase in genome instability recombination events in RNase H2-deficient pol2-M644G cells. This is to point out that there are multiple ways to process ribonucleotides in DNA that lead to recombination-initiating events, and indeed not all may be DSBs, but may be Top1-dependent.

Response: We thank the reviewer for noting the complexity in interpreting the in vivo data, and in response we have addressed these issues in the text. The O'Connell et al. and Epshtein et al. papers are now cited in the Introduction. We include a description of the fact that the Conover paper reported increased LOH both genome-wide (after passaging diploid yeast strains) and in their

particular LOH assay. Although the accompanying O'Connell paper came to the conclusion that RNA-DNA hybrids are the primary source of recombination in RNase H-deficient yeast, they did not present results for the *pol2-M644G rnh201*D strain. However, their conclusion is based on failure to observe a decrease in LOH using the pol2-M644L variant of polymerase epsilon that incorporates fewer ribonucleotides than the wild-type enzyme (3-fold fewer versus 10-fold more for the pol2-M644G allele, as measured in vitro). We now also discuss Epshtein et al., 2016, where they did not observe elevated recombination in the *pol2-M644G rnh201*D strain using their reporter system. We agree that there are multiple ways that Top1-processing of ribonucleotides can lead to recombination and we made an effort to clarify this point in the Discussion.

Figure 2a shows that top1 mutation rescues viability of pol2-M644G rnh201 cells, but viability is far from normal. This needs some discussion. Does this mean that there are still breaks but above the threshold for no cells dividing? Or are breaks repaired but in a manner that gives rise to rearranged genomes? Or is this related to the general malaise of the pol2-M644G strain? Similarly, the pol1 and pol3 mutations in the rnh201 strain are not normal. This could suggest that processed or tolerated ribonucleotides are genome-destabilizing in some manner.

Response: We thank the reviewer for raising this point. We have now included in the Discussion a description of the fact that although viable, the pol2-M644G rnh201D rad52D top1D strain is slow growing and that this may reflect DNA break formation and genome rearrangements. Although this may be related to the *pol2-M644G* allele, a *pol2-M644G rnh201*D *top1*D haploid strain grows remarkably well and does not display checkpoint activation or replication stress above that of the *pol2-M644G* single mutant (Williams, 2013) suggesting that the slow growth of the *pol2-M644G rnh201*D rad52D *top1*D strain has more to do with loss of DSB recombination repair than any general malaise associated with the *pol2-M644G* variant. We agree that the *pol1-LM rnh201*D *rad52*D and *pol3-LM rnh201*D *rad52*D strains are also slow growing and that this may reflect ribonucleotide-dependent genome stability. This point is now included in the Discussion.

I would also like to see an expanded discussion of the model in Figure 4. If the ribonucleotide is in a repeat sequence, then as has been shown in vivo and in vitro, a second Top1 cleavage near the cyclic phosphate and slippage can result in a deletion without a DSB being formed. How the second cleavage and realignment are coordinated is not entirely understood. However, if the ribonucleotide is not in a repeat, which may be more frequent, there may be no slippage realignment so a cleavage could occur on the same strand as the cyclic phosphate or on the opposite strand to give a DSB. Religation and processing (the cyclic phosphate removal in the left of the f panel is not addressed) could give some type of gap filling or recombination. Based on the ribonucleotide misincorporation rate and the reported recombination rates (which do not detect all recombination events), it may be that religation/gap filling is more prevalent.

Response: We thank the reviewer for the careful analysis of the published data on recombination rates and for raising this interesting point. We have now included in the manuscript an expanded discussion on previous studies that elude to possible DNA gap-filling repair for Top1-induced nicks at newly-incorporated ribonucleotides. We have also included an expanded figure regarding the model in Figure 4. It is now Fig EV5. It summarizes the multiple outcomes and the biological consequences of incorporated ribonucleotide processed by Top1.

Minor note:
Some of the data are redundant, such as tables and graphs of the same data set, especially in the supplementary material (eg graphs and tables).

Response: Thank you. We have consolidated the data previously shown in 3 separate tables into Appendix Table S1 and Figure EV2 in the current revision.

Referee #3:

Overall, the manuscript has improved a lot. Results are much stronger.
Beautiful work.

Response: We thank the reviewer for qualifying our study as "beautiful work". Thank you.

3rd Editorial Decision 4 November 2016

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

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Yves Pommier
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2015-92426R

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

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

2. Captions

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

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

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

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

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

In the pink boxes below, provide the page number(s) of the manuscript draft or figure legend(s) where the information can be located. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

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?	For the spontaneous Rad52-YFP foci measurement, at least 3200 cells for each strain was counted to ensure statistical power. The cell cycle distribution and growth rates was meaured independently at least twice.
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?	For the analysis of DNA covalently-linked with Top1, we only proceed with S1-treatment when the DNA from WT and RNase H2-null samples were comparable.
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.	NA
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, please see relevant figure legends for description of every statistical test performed.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	NA
Is there an estimate of variation within each group of data?	Yes, either SD or 95% confidence interval is presented.
Is the variance similar between the groups that are being statistically compared?	Yes.

C- Reagents

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-tum

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://jij.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	See relevant section in Materials and Methods and Figure Legends.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	See the Yeast Strains section in Materials and Methods.

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

D- Animal Models

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

E- Human Subjects

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

F- Data Accessibility

18. Provide accession codes for deposited data. See author guidelines, under '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. As far as possible, primary and referenced data should be formally cited in a Data Availability section. Please state whether you have included this section. Examples: **Primary Data** Wetmore KM, Deutschbauer AM, Price MN, Arkin AP (2012). Comparison of gene expression and mutant fitness in Shewanella oneidensis MR-1. Gene Expression Omnibus GSE39462 **Referenced Data** Huang J, Brown AF, Lei M (2012). Crystal structure of the TRBD domain of TERT and the CR4/5 of TR. Protein Data Bank 4O26 AP-MS analysis of human histone deacetylase interactions in CEM-T cells (2013). PRIDE PXD000208	NA
22. 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

23. 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