nature research

Peer Review Information

Journal: Nature Cell Biology
Manuscript Title: The PPR domain of mitochondrial RNA polymerase is an exoribonuclease required for mtDNA replication in *D. melanogaster*
Corresponding author name(s): Dr Hong Xu

Reviewer Comments & Decisions:

Decision Letter, initial version:

Subject: Decision on Nature Cell Biology submission NCB-X45730
Message:

*Please delete the link to your author homepage if you wish to forward this email to co-authors.

Dear Dr Xu,

Your manuscript, "The PPR domain of mitochondrial RNA polymerase is a ribonuclease required for mtDNA replication", has now been seen by 3 referees, who are experts in mitochondrial DNA and Drosophila (referee 1) and mitochondrial DNA (referees 2 and 3). As you will see from their comments (attached below) they find this work of potential interest, but have raised substantial concerns, which in our view would need to be addressed with considerable revisions before we can consider publication in Nature Cell Biology.

Nature Cell Biology editors discuss the referee reports in detail within the editorial team, including the chief editor, to identify key referee points that should be addressed with priority, and requests that are overruled as being beyond the scope of the current study. To guide the scope of the revisions, I have listed these points below. We are committed to providing a fair and constructive peer-review process, so please feel free to contact me if you would like to discuss any of the referee comments further.

In particular, it would be essential to:

a) use proper substrates that are relevant for in vivo primer formation in the nuclease activity assays, as noted by:

Referee 1:
As far as I can see the nuclease activity of the isolated PPR domain was tested only on a ssRNA substrate, and the assays should therefore be repeated (using the isolated PPR domain and E423P mutant, plus the full-length protein and mutant) on an RNA/DNA hybrid substrate. Understanding the interaction with hybrid may be crucial to understanding the role of the PPR domain in vivo: I would predict that the findings shown in Fig. 3, i.e. the ability of the enzyme to prime DNA synthesis by Poly, are best explained if the PPR domain both degrades ssRNA and stabilizes otherwise transient RNA/DNA hybrid, but does not actually degrade hybrid. In other words creating an RNA 3´ end at a site of hybrid formation. This was not formally demonstrated and is an easy experiment to perform. The final experiments of the paper, showing that the nuclease activity of the PPR domain is required to remove erroneously incorporated bases in mtRNA, raises the intriguing possibility that the PPR domain functions in mismatch surveillance, activating the exonuclease activity when a mismatched base-pair is detected in RNA/DNA hybrid, as the transcription complex advances. As an initial test of this, it would therefore also be worth testing mismatched RNA/DNA substrates in EMSA for binding and in nuclease assays for trimming activity. Without such experiments, mechanistic understanding of the role of the PPR domain will remain rather minimal.

Referee 3:
2. The exclusive use of single-stranded M13 DNA at the template for all in vitro transcription analyses is a limitation of the study. The in vivo substrate for primer formation is double-stranded mtDNA and, in the case of the leading-strand origin, primers are generated from RNA derived from promoter-specific transcription events. Can the authors show with either the Drosophila or human POLRMT that the exonuclease activity is relevant for primer formation in this more relevant context? This would also allow a better conclusion that the PPR version of the enzyme has normal transcription activity.

b) provide missing controls in relevant experiments, as noted by:

Referee 1:
b/ Nucleic-acid binding and nuclease assays: the statement in lines 5-7 on p. 6 ("Given that PPR motifs have been proposed to be RNA binding proteins, we hypothesized that PPR might bind to some short DNA sequences to initiate RNA synthesis from these discrete sites on mtDNA") seems fundamentally illogical. If it binds to RNA, why should we expect it to bind also to (ds?)DNA. The experiment of Fig. 2 is nevertheless justified, just not on this basis. In Fig. 2 the authors show data indicating that the PPR domain of PolrMT binds to RNA/DNA hybrid, and may degrade hybrid as well as ssRNA exonucleolytically. This experiment is crucial to the paper, but there are some missing controls. First, the influence of contaminating E. coli proteins has not been rigorously excluded. Data on the purity and integrity of the isolated PPR domain should be more thoroughly presented in Fig. S3 (e.g. using silver-

stained SDS-PAGE gels to document the purification steps and final purity). A protein isolate from control cells bearing the expression vector only, treated identically as the experimental isolate, needs to be tested in parallel. Most importantly, all of the EMSA and nuclease assays should be repeated with the isolated PPR domain bearing the E423P 'nuclease-deficient' mutation – as far as I can see only the full-length protein bearing this mutation was tested in the experiment of Fig. 3. This control is especially important since the data of Fig. 2A imply that the PPR domain of PolrMT might have an RNase H-like activity, potentially making it a second RNase H inside mitochondria.

Referee 2:
Figure 4 – In order to determine whether the POLMRT restored the steady-state levels of mtDNA and mitochondrial transcripts, wild type control fly strain results should be shown for all data. Otherwise, it is unclear whether the overexpression of transgenic POLMRT is increasing mtDNA and mt-RNA levels above wild type, whereas POLMRTE423P does not.

c) strengthen the data to support the PPR domain of POLRMT as the exoribonuclease to synthesize the primer, as noted by Referee 3:

1. From the results in Figure 1D, the authors conclude that the main effect of deleting the PPR domains (lane 4, ΔPPR) is inability to synthesize shorter RNA transcripts. However, there is very little signal here at all, suggesting an effect on RNA polymerization as well. Thus, it is difficult to conclude that the ΔPPR enzyme is only deficient in synthesizing short RNA transcripts.

3. While the data are largely consistent with the PPR domain of POLRMT possessing 3-prime exonuclease activity, the gels demonstrating this directly are somewhat atypical of a processive exonuclease. That is, there is an expected progressive loss of full-length products over time, but the smaller degradation products do not gradually increase in concert in Figures 2B, 2C and 3A. In other words, the full array of expected degradation products that differ by one nucleotide is observed at the very first timepoint (5 minutes) and do not steadily increase or shift to smaller size ranges over time as expected. How do the authors explain this? Also, a negative control including the reaction mix and the substrate, but not the enzyme, subjected to the longest incubation period in the experiment (30 mins) should be included.

d) All other referee concerns pertaining to strengthening existing data, providing controls, methodological details, clarifications and textual changes, should also be addressed.

e) Finally please pay close attention to our guidelines on statistical and methodological reporting (listed below) as failure to do so may delay the reconsideration of the revised manuscript. In particular please provide:

- a Supplementary Figure including unprocessed images of all gels/blots in the form of a multi-page pdf file. Please ensure that blots/gels are labeled and the sections presented in the figures are clearly indicated.

- a Supplementary Table including all numerical source data in Excel format, with data for different figures provided as different sheets within a single Excel file. The file should include source data giving rise to graphical representations and statistical descriptions in the paper and for all instances where the figures present representative experiments of multiple independent repeats, the source data of all repeats should be provided.

We would be happy to consider a revised manuscript that would satisfactorily address these points, unless a similar paper is published elsewhere, or is accepted for publication in Nature Cell Biology in the meantime.

When revising the manuscript please:

- ensure that it conforms to our format instructions and publication policies (see below and www.nature.com/nature/authors/).

- provide a point-by-point rebuttal to the full referee reports verbatim, as provided at the end of this letter.

- provide the completed Editorial Policy Checklist (found here https://www.nature.com/authors/policies/Policy.pdf), and Reporting Summary (found here https://www.nature.com/authors/policies/ReportingSummary.pdf). This is essential for reconsideration of the manuscript and these documents will be available to editors and referees in the event of peer review. For more information see http://www.nature.com/authors/policies/availability.html or contact me.

Nature Cell Biology is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit www.springernature.com/orcid.

Please submit the revised manuscript files and the point-by-point rebuttal to the referee comments using this link:

[REDACTED]

*This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

We would like to receive a revised submission within six months. We would be happy to consider a revision even after this timeframe, however if the resubmission deadline is missed and the paper is eventually published, the submission date will be the date when the revised manuscript was received.

We hope that you will find our referees' comments, and editorial guidance helpful. Please do not hesitate to contact me if there is anything you would like to discuss.

Best wishes,

Jie Wang

Jie Wang, PhD
Senior Editor
Nature Cell Biology

Tel: +44 (0) 207 843 4924
email: jie.wang@nature.com

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
This paper reports the characterization of the PPR domain of the Drosophila mitochondrial RNA polymerase. The domain was found to be required for the synthesis of short, as opposed to long, transcripts on a ssDNA template, conferred binding to RNA/DNA hybrid and had 3´ to 5´ exoribonuclease activity. In an artificial in vitro system, the active exonuclease was needed to enable DNA synthesis, and was also needed for mtDNA replication during oogenesis in vivo. Sequence analysis of mitochondrial

RNA indicated that the exonuclease activity of the PPR domain was needed for proofreading during transcription.

Although I have many issues concerning missing controls and data presentation, the main substance of the paper seems to be sound, and its findings are an original and important contribution to the field.

In revising the manuscript, the authors should pay attention to the following points. I apologize for their length, but I feel it is important to be specific and detailed as to what needs to be done to get the manuscript into shape, and why.

1. Nomenclature

Although the manuscript is mainly about Drosophila enzymes, the authors use the human nomenclature almost throughout. Flybase is encouraging authors to use official gene and protein names, so I strongly suggest that at first mention of each protein both the mammalian and fly names are given, e.g. mitochondrial RNA polymerase (human – POLRMT, mouse – Polrmt, Drosophila – PolrMT, CG4644) and thereafter use the species-specific nomenclature as appropriate, at each mention.

2. Other definitions

Throughout the manuscript the authors do not distinguish between leading- and lagging-strand priming. Although little is known of either process in Drosophila, they are likely to be fundamentally different, as in almost all other biological contexts, with leading-strand priming almost certainly requiring transcription from a locally unwound dsDNA template and lagging-strand priming taking place on displaced ssDNA either at the fork or elsewhere. Since almost all of the priming-related experiments presented here used a ssDNA substrate, it should be made clearer that the PPR domain of PolrMT is here implicated in lagging-strand priming, with the mechanics of leading-strand priming left open.

3. Experimental findings

a/ Deletion phenotype (Fig 1): larvae with a clean deletion of PolrMT were small (1A), arrested at L3, and showed no puncta of mtDNA synthesis in fat body (1B). The purified protein was active in synthesizing short RNAs on an M13 ssDNA template whereas T7 pol and the ΔPPR variant synthesized mostly long RNAs and the ΔPPR variant had only a very low activity although the latter point is glossed over. Inclusion of the missing ΔNTE-PPR construct in the assay is really needed to address the issue of whether

the PPR domain is needed for full polymerase activity as well as to limit the production of long transcripts.

b/ Nucleic-acid binding and nuclease assays: the statement in lines 5-7 on p. 6 ("Given that PPR motifs have been proposed to be RNA binding proteins, we hypothesized that PPR might bind to some short DNA sequences to initiate RNA synthesis from these discrete sites on mtDNA") seems fundamentally illogical. If it binds to RNA, why should we expect it to bind also to (ds?)DNA. The experiment of Fig. 2 is nevertheless justified, just not on this basis. In Fig. 2 the authors show data indicating that the PPR domain of PolrMT binds to RNA/DNA hybrid, and may degrade hybrid as well as ssRNA exonucleolytically. This experiment is crucial to the paper, but there are some missing controls. First, the influence of contaminating E. coli proteins has not been rigorously excluded. Data on the purity and integrity of the isolated PPR domain should be more thoroughly presented in Fig. S3 (e.g. using silver-stained SDS-PAGE gels to document the purification steps and final purity). A protein isolate from control cells bearing the expression vector only, treated identically as the experimental isolate, needs to be tested in parallel. Most importantly, all of the EMSA and nuclease assays should be repeated with the isolated PPR domain bearing the E423P 'nuclease-deficient' mutation – as far as I can see only the full-length protein bearing this mutation was tested in the experiment of Fig. 3. This control is especially important since the data of Fig. 2A imply that the PPR domain of PolrMT might have an RNase H-like activity, potentially making it a second RNase H inside mitochondria. As far as I can see the nuclease activity of the isolated PPR domain was tested only on a ssRNA substrate, and the assays should therefore be repeated (using the isolated PPR domain and E423P mutant, plus the full-length protein and mutant) on an RNA/DNA hybrid substrate. Understanding the interaction with hybrid may be crucial to understanding the role of the PPR domain in vivo: I would predict that the findings shown in Fig. 3, i.e. the ability of the enzyme to prime DNA synthesis by Polγ, are best explained if the PPR domain both degrades ssRNA and stabilizes otherwise transient RNA/DNA hybrid, but does not actually degrade hybrid. In other words creating an RNA 3´ end at a site of hybrid formation. This was not formally demonstrated and is an easy experiment to perform. The final experiments of the paper, showing that the nuclease activity of the PPR domain is required to remove erroneously incorporated bases in mtRNA, raises the intriguing possibility that the PPR domain functions in mismatch surveillance, activating the exonuclease activity when a mismatched base-pair is detected in RNA/DNA hybrid, as the transcription complex advances. As an initial test of this, it would therefore also be worth testing mismatched RNA/DNA substrates in EMSA for binding and in nuclease assays for trimming activity. Without such experiments, mechanistic understanding of the role of the PPR domain will remain rather minimal.

c/ Regulation of nuclease activity: the exonuclease activity of the PPR domain has to be context-regulated in some way, or else no primer RNA (or any RNA) would ever be made. The nuclease activity should therefore also be measured using a variant of the full-length protein, incorporating a ‘polymerase-dead’ mutation. The issue of how the exonuclease activity is activated or limited is referred

to at the foot of p. 11/top of p. 12, but in a confusing way, since TEFM appears to have no orthologue in Drosophila, so its involvement is irrelevant. Conversely, the experiments presented here relate to the Drosophila enzyme, and the properties of the PPR domain of the mammalian enzyme may differ. I would suggest to make these points clearer. See also next point, 3d.

d/ Role in synthesis of short DNA fragments: overexpression in vivo of the wild-type enzyme but not the E423P mutant restored viability and mtDNA maintenance, and supported the production of short mtDNA fragments in organello (Fig. 4C). The authors comment that this shows an intriguing similarity with the synthesis of 7S DNA in mammalian mitochondria. However, since there is no D-loop in Drosophila mtDNA, no previous detection of an abundant short DNA strand, and the molecules synthesized in organello were not in any way mapped, I think this assertion is fanciful and should be dropped. The fact that only short DNA fragments were synthesized suggests that some other component required for processive mtDNA replication is missing or disabled in the in organello system used.

e/ Adult phenotype: the experiments on adult phenotype are, indeed, intriguing, although the authors' statement that they were puzzled by the findings is a little misleading, since the subsequent experiments indicate a likely explanation. Maybe it would suffice just to say that the results were 'initially' puzzling. One other small issue concerns the statement that there is little mtDNA replication in the adult musculature. The measurements of copy number reported in Fig. S4 are all indirect: whilst suggesting that the expression of the mutant enzyme has no effect on copy number, it's not very clear why the authors didn't simply measure copy number in isolated thoraxes of males and females of the different genotypes by qPCR, which would have settled the issue more convincingly. Another variable that might have compromised the data of Fig. 5 concerns the expression levels of the inserted transgenes. Western blot data in Fig. S4C probably answers this but needs proper annotation (see point 6q, below). The targeted insertion sites (plus confirmation that these were the (only) insertions detected in the lines) should be indicated somewhere, e.g. in 'Methods', and qRT-PCR from different tissue and stages would help convince me that the mutant phenotype is due to the mutation per se, rather than over or mis-expression.

5. Methods

a/ Substrates: the authors state that "all 6-FAM and biotin labelled RNA oligos including oligos used as markers were synthesized by IDT DNA" – what is IDT DNA? – I assume it is a company, but please give company name and location in full. The creation of RNA/DNA hybrid substrates for EMSA is not described.

b/ Climbing assay: it isn't clear to me how the final scores were computed. Please be more specific.

c/ S2 cells: The 'Methods' incudes a section on culture conditions for S2 cells – however, as far as I could see, no use was made in the paper of S2 cells.

6. Data quality, figures and statistics

a/ Fig. 1A – scale bars should be shown, and the mt and mutant larval mouthparts should be shown to the same scale in the rh panel

b/ Fig. 1B – is anecdotal. Quantitative data should be added: how many animals of each genotype were analysed, how many TFAM puncta were analysed, how many were EdU positive. Are wt and mutant larvae of the same chronological or developmental age, and at what stage (L3 or something else?). Dissection details or a reference should be added to 'Methods'. Why was this experiment not done also on larvae expressing the E423P variant? This is an important experiment to test whether the ribonuclease activity of PolrMT is needed for mtDNA synthesis in vivo.

c/ Fig. 1C legend – please make clear if/that the numbering refers to the Drosophila enzyme PolrMT. No data are presented on the purity and intactness of the expressed proteins (e.g. by SDS-PAGE), which should be shown in a supplementary figure: standard procedure for validating the data shown in Fig. 1D.

d/ Fig. 1D – once again, only qualitative data from a single experiment are shown. The track for the ΔPPR variant is so faint, whilst that for ΔNTE so black/overexposed, no reliable conclusion can be drawn regarding the role of PPR domain is favouring short RNA products. Some quantitation by phosphorimaging should be added, a longer exposure of the ΔPPR track should be shown, and markers in the size-range of replication primers should be added. Finally, one informative construct is missing, ΔNTE-PPR. and should be included.

e/ Fig. 3D – the legend (and text) do not make clear whether the assay used fly or human Polγ.
I can see no sense in combining the fly PolrMT and human Polγ, but the isolation of the fly Polγ is not described anywhere. And was the PPR domain used in this experiment in fact from the human RNA polymerase. Or the fly enzyme? Since the paper is primarily about the fly enzymes, it is extremely confusing to include experiments on the human enzymes without clearly stating that this is the case and justifying the switch of species. The authors should clarify and justify the provenance of the enzymes used here and in each experiment throughout the paper. Given the absence of two relevant accessory proteins in Drosophila (TEFM and Primpol), experiments using the human enzymes are not, in my opinion, really justified here, except as supplementary data to show that the human system behaves similarly (or not) to the Drosophila one (e.g. as in Fig. S3). Note also that Polγ from the two species even has a different subunit composition, and so a chimeric experiment using Polγ from one species and

PolrMt (or its PPR domain) from the other is simply not reliable in this context. In short, the main experiment should be repeated using Drosophila Polγ.

f/ Fig. 4A – would make more sense if mtRNA levels were normalized to mtDNA. The wt and E423P mutants would then give rather similar transcript levels, while both would be a bit higher than the deletion (though the data might not be statistically significant). Please specify how 'relative level' was normalized: to what? Wild-type?

g/ Fig. 4B – the red signal in the top rh panel (arrowed) is almost invisible. Please supply higher quality images.

h/ Fig. 4C - 'relative level' of what, normalized to what? At least refer the reader back to 'Methods' where this seems to be specified.

i/ Fig. 4A, C – I found the description of the genotypes sloppy and confusingly presented, especially with the inserted commas. Please use standard Flybase symbols, specify the allele of w, clarify whether ac-Gal4 (? Act5C or achaete?) and UAS-PolrMT are really both on chromosome 3 and remove duplication of PolrMT[del] where flies are homozygous. Specify accurately the source of all lines in 'Methods' or in a supplementary table.

j/ Fig. 5B – although it's in 'Methods', please add to the legend that these were male flies

k/ Fig. 5C – please clarify what is meant in the legend by 'time of a misfit'

l/ Fig. 5D – the pictures are fine as an example, but are of such poor resolution that indicating mt by arrows is pointless. However, the authors should define 'selected mitochondrial area' more precisely in 'Methods'.

m/ Fig. 5 F-H – why are wild-type control flies not shown in these experiments? Please determine and add the values for wt controls.

n/ Fig. 5F – this is a busy figure panel which says rater little. I would think it sufficient to appear in the supplementary figures, especially as the functional significance of poly(A) length in mitochondrial transcripts is not properly understood.

o/ Fig. 5 legend – once again, the genotypes are confusingly presented. Please use Flybase-approved official symbols and add the information on the Gene Switch strain (again using approved nomenclature) to 'Methods'.

p/ Fig. S4A – the genotype of the flies is unstated and no controls are shown.

q/ Fig. S4C – the 'POLRMT Bac clone' line is not described anywhere else in the paper. Its construction must be described and justified for use in this experiment. Please also specify where proteins were extracted from: L3 larvae, whole adults (male/female?), thoraxes?

r/ Fig. S5 title – is misleading, since the very next figure shows that it isn't 'unaffected'. I suggest instead 'The ablation of the exoribonuclease activity of PolrMT does not affect the amount of mitochondrial transcription'.

7/ References

No major comments, but some rearrangement and rebalancing should be needed, to rectify the problems with the Introduction (see below).

8. Introduction

Although the PPR domain is found in the mitochondrial RNA polymerase from many species, the specific context of Drosophila in the present experiment needs to be more strongly emphasized and justified. Importantly, the replication mechanism of mtDNA in Drosophila is not known in great detail, but appears to differ from the vertebrate one,. This needs to be made clearer and the experiments to be undertaken in the paper should be cast as being aimed at understanding what is happening in the fly model, rather than directly extrapolating seamlessly to the human/vertebrate context. The very first paragraph introduces 'mammalian mtDNA', but instead should first describe animal mtDNA in general, then more specifically Drosophila mtDNA, and finally data that has been obtained in the mammalian context, for comparison. In all cases it should be made clear what is general and what is or may be unique to mammals (or flies). As currently written, the bulk of the section is about mammalian mtDNA without actually saying so, and this is both irritating and misleading to the reader. Note also that evidence for the strand displacement replication model across most of the mitochondrial genome of Drosophila is weak, whilst strand-coupled replication is well documented. The statement about Primpol is nonsensical: since it is found only in chordates, as stated, it makes no sense to explore its prevalence and potential roles in other animal phyla, where it is absent!

9/ Minor typos

p. 1, line 22 – 'transcriptional', not 'transcripts'

p. 10, line 26 - ‘correct’, not ‘corrected’

Reviewer #2:
Remarks to the Author:
Liu et al report on the characterisation of the PPR domain of Drosophila POLMRT as a primase for mtDNA replication and proof-reader for mitochondrial transcription, though its exoribonuclease activity. The authors used a wide range of biochemical experiments to demonstrates these actions and further determined the consequences of disrupting these activities on fly physiology and behaviour. The study is well controlled in most cases and the authors demonstrated awareness of possible confounding factors to their interpretations that were then carefully excluded. Overall, this is a very nice study that is important for the field.

Comments and Questions

The authors have found that fly POLMRT acts in a similar manner to mammalian POLMRT.
Figure 4 – In order to determine whether the POLMRT restored the steady-state levels of mtDNA and mitochondrial transcripts, wild type control fly strain results should be shown for all data. Otherwise, it is unclear whether the overexpression of transgenic POLMRT is increasing mtDNA and mt-RNA levels above wild type, whereas POLMRTE423P does not.

The position of the PPR domain shown in Figure 1C is different to uniport predictions. Could the authors specify the source of domain prediction?

In figure 3C, the intensity of biotin ssRNA level between POLRMTE423P and T7 RNA polymerase are quite similar, however in the line chart next to it, why is the level of biotin-UTP integration of POLRMTE423P much lower than T7 RNA polymerase over the time. Could the authors please explain this difference?

In figure S5, the authors claimed that ablation of POLRMT’s exoribonuclease activity does not affect mitochondrial transcription. However, in figure 4, the expression of POLRMTE423P in POLRMTdel mutant was unable to restore the mitochondrial transcription. Could the authors please explain this opposite result?

Minor points
The article has many grammatical errors and would require proofreading.

Figure 1, 4, Figure S3, S4 panel order does not match textual order.

P. 7 line 27 and 29 – There is no Figure S3F

Consider a heading for the discussion section of the manuscript.

Reviewer #3:
Remarks to the Author:
The mechanisms underlying the control of animal mtDNA replication remain to be fully elucidated and important to understand not only from a fundamental perspective, but also because defects in this process lead to human diseases. There is a general consensus that mtDNA replication in mammals (studied mostly in humans and mice) is primed by RNA transcripts synthesized by the mtRNA polymerase (POLRMT), but how short primer RNAs are generated in addition to "full-length" genome-sized transcripts has been a controversial area for many years. For the heavy-strand origin a nuclease, RNAse MRP, was proposed to process POLRMT transcripts via an endonuclease activity in the context of a complex RNA-DNA hybrid to generate RNA primers. However, others have since proposed that regulated transcription termination to form a complex RNA/DNA hybrid that is processed by RNAseH1 instead generates the primers. For the light-strand origin, a similar history exists. A primase enzyme was originally proposed, but was not able to be identified, and the Falkenberg group subsequently proposed that POLRMT is also the OL primase, but with no proposed requirement for a nuclease to generate the primers. The current study by Liu et al. is significant because they provide compelling evidence that POLRMT harbors an exonuclease activity that enables it to generate short RNA primers for mtDNA replication. Specifically, they provide evidence that the PPR domains of Drosophila and human POLRMT have intrinsic 3'-to-5' exonuclease activity. In vitro experiments with a truncated form of POLRMT deficient in exonuclease activity (ΔPOLRMT) indicate that this nuclease activity is required for primer formation, but not RNA polymerization activity (i.e. transcription). In flies, in vivo overexpression of a mutant POLRMT lacking only the exonuclease activity (POLRMTE423P) results in an increase in mtDNA transcription errors and neuromuscular pathology, implying another role of the ribonuclease activity of POLRMT in proofreading to increase transcriptional fidelity. This is an interesting and important study with high relevance to our understanding of mtDNA replication and elucidates two novel functions of POLRMT, nuclease-mediated primer formation and transcriptional fidelity. There are however some concerns that need to be addressed in order bolster the conclusions reached by the authors that are detailed below:

Main points:

1. From the results in Figure 1D, the authors conclude that the main effect of deleting the PPR domains (lane 4, PPR) is inability to synthesize shorter RNA transcripts. However, there is very little signal here

at all, suggesting an effect on RNA polymerization as well. Thus, it is difficult to conclude that the PPR enzyme is only deficient in synthesizing short RNA transcripts.

2. The exclusive use of single-stranded M13 DNA at the template for all in vitro transcription analyses is a limitation of the study. The in vivo substrate for primer formation is double-stranded mtDNA and, in the case of the leading-strand origin, primers are generated from RNA derived from promoter-specific transcription events. Can the authors show with either the Drosophila or human POLRMT that the exonuclease activity is relevant for primer formation in this more relevant context? This would also allow a better conclusion that the PPR version of the enzyme has normal transcription activity.

3. While the data are largely consistent with the PPR domain of POLRMT possessing 3-prime exonuclease activity, the gels demonstrating this directly are somewhat atypical of a processive exonuclease. That is, there is an expected progressive loss of full-length products over time, but the smaller degradation products do not gradually increase in concert in Figures 2B, 2C and 3A. In other words, the full array of expected degradation products that differ by one nucleotide is observed at the very first timepoint (5 minutes) and do not steadily increase or shift to smaller size ranges over time as expected. How do the authors explain this? Also, a negative control including the reaction mix and the substrate, but not the enzyme, subjected to the longest incubation period in the experiment (30 mins) should be included.

4. The authors need to do a better job of reconciling and contextualizing their results with prior published results on mtDNA replication priming. For example, others have implicated transcription termination and RNAse H1 processing of a complex RNA/DNA hybrid in priming OH in mammals and no nuclease has been implicated to prime POLRMT-derived transcripts OL. There is also evidence for direct role for mtSSB in priming that should be discussed. In addition, are the results on deleting the NTE (Fig. 1D) consistent with other published findings that have deleted this region of POLRMT (Posse et al. NAR 2014 42:3638)? Finally, any salient differences between mammalian mtDNA replication priming and Drosophila should be highlighted more. The same should be done for mtRNA processing if relevant. For example, the statement on Page 3, lines 4–5, "all mitochondrial mRNAs and rRNAs are polyadenylated", citing Mercer et al. (2011), is incorrect. According to that study, ND6 transcript is not polyadenylated in humans. Is this true in fly as well? Please clarify.

5. Wang and Shadel 1999 (PNAS 96:8046) performed N-terminal deletion analysis of yeast mtRNA polymerase (RPO41) and showed it is required for mtDNA stability. Interestingly, they also showed that one deleted version, while fully capable of RNA polymerization, was no longer capable of generating a prominent shorter transcript from an origin-containing template in vivo (see Fig. 3C). Kruszewski and Golik (Biochemistry/Moscow 2016 81:1101) subsequently showed that this deleted region contains two degenerate PPR domains, suggesting that a PPR-associated nuclease activity of POLRMT may be conserved in yeast. These results seem consistent with the current findings and should be discussed.

6. Please provide more detail on the bioinformatic analysis of unprocessed transcripts as this is not a common practice and could be useful to others in the field. For example, specific questions include the following: Which junctions were included? For example, was the Atp8-Atp6 junction included, which encodes parts of both genes and hence likely left uncleaved during mRNA maturation? What was the criteria chosen to call a read "gene spanning"? Is including 1 nucleotide on gene A and 50 nucleotides on a neighboring gene B sufficient to make a call? How was SAMtools configured to count gene-spanning reads instead of reads mapping to intragenic regions? Was any input (a custom list of sequences, filtering criteria, etc.) provided by the authors to SAMtools? Were any parameters modified? Lastly, please provide the number of flies used to isolate mitochondria for RNA sequencing and the number of samples that were sequenced independently (indicate any sample pooling that was used).

Minor points.

• Please indicate the age of larva used in the EdU incorporation assay shown in Figure 1B either in the main text or in the methods section.

• In Figure 1B, if possible, it might be better to show another area on POLRMTdel sample that includes at least one nucleus so that EdU incorporation into the nuclear DNA is shown while it is lacking in the mitochondria, i.e., the lack of EdU signal is not due to a total failure of DNA replication (both nuclear and organellar) at the larval stage tested in this lethal POLRMT deletion model.

• Please indicate the antibody used to stain TFAM.

• In the methods, a normal "x" letter should be used, instead of the multiplication symbol "×", to indicate the concentration of solutions. For example, 10x lysis buffer, 2x YT medium, etc.

• Some arrowheads described in the figure legends are missing in the figures. See figures 3A, S2C, S2D, S3E and corresponding figure legends, for example.

• Since the majority of the results of the study are with Drosophila and the replication function was only demonstrated in flies, should the title be edited as follows (or in some other way) to reflect this? “The PPR domain of mitochondrial RNA polymerase is a ribonuclease required for mtDNA replication in D. melanogaster”.

• Will the mitochondrial RNA sequencing data be made available online?

GUIDELINES FOR MANUSCRIPT SUBMISSION TO NATURE CELL BIOLOGY

READABILITY OF MANUSCRIPTS – Nature Cell Biology is read by cell biologists from diverse backgrounds, many of whom are not native English speakers. Authors should aim to communicate their findings clearly, explaining technical jargon that might be unfamiliar to non-specialists, and avoiding non-standard abbreviations. Titles and abstracts should concisely communicate the main findings of the study, and the background, rationale, results and conclusions should be clearly explained in the manuscript in a manner accessible to a broad cell biology audience. Nature Cell Biology uses British spelling.

MANUSCRIPT FORMAT – please follow the guidelines listed in our Guide to Authors regarding manuscript formats at Nature Cell Biology.

TITLE – should be no more than 100 characters including spaces, without punctuation and avoiding technical terms, abbreviations, and active verbs..

AUTHOR NAMES – should be given in full.

AUTHOR AFFILIATIONS – should be denoted with numerical superscripts (not symbols) preceding the names. Full addresses should be included, with US states in full and providing zip/post codes. The corresponding author is denoted by: "Correspondence should be addressed to [initials]."

ABSTRACT AND MAIN TEXT – please follow the guidelines that are specific to the format of your manuscript, as listed in our Guide to Authors (http://www.nature.com/ncb/pdf/ncb_gta.pdf) Briefly, Nature Cell Biology Articles, Resources and Technical Reports have 3500 words, including a 150 word abstract, and the main text is subdivided in Introduction, Results, and Discussion sections. Nature Cell Biology Letters have up to 2500 words, including a 180 word introductory paragraph (abstract), and the text is not subdivided in sections.

ACKNOWLEDGEMENTS – should be kept brief. Professional titles and affiliations are unnecessary. Grant numbers can be listed.

AUTHOR CONTRIBUTIONS – must be included after the Acknowledgements, detailing the contributions of each author to the paper (e.g. experimental work, project planning, data analysis etc.). Each author should be listed by his/her initials.

FINANCIAL AND NON-FINANCIAL COMPETING INTERESTS – the authors must include one of three declarations: (1) that they have no financial and non-financial competing interests; (2) that they have financial and non-financial competing interests; or (3) that they decline to respond, after the Author Contributions section. This statement will be published with the article, and in cases where financial and non-financial competing interests are declared, these will be itemized in a web supplement to the article. For further details please see https://www.nature.com/licenceforms/nrg/competing-interests.pdf.

REFERENCES – are limited to a total of 70 for Articles, Resources, Technical Reports; and 40 for Letters. This includes references in the main text and Methods combined. References must be numbered sequentially as they appear in the main text, tables and figure legends and Methods and must follow the precise style of Nature Cell Biology references. References only cited in the Methods should be numbered consecutively following the last reference cited in the main text. References only associated with Supplementary Information (e.g. in supplementary legends) do not count toward the total reference limit and do not need to be cited in numerical continuity with references in the main text. Only published papers can be cited, and each publication cited should be included in the numbered reference list, which should include the manuscript titles. Footnotes are not permitted.

METHODS – Nature Cell Biology publishes methods online. The methods section should be provided as a separate Word document, which will be copyedited and appended to the manuscript PDF, and incorporated within the HTML format of the paper.

Methods should be written concisely, but should contain all elements necessary to allow interpretation and replication of the results. As a guideline, Methods sections typically do not exceed 3,000 words. The Methods should be divided into subsections listing reagents and techniques. When citing previous methods, accurate references should be provided and any alterations should be noted. Information must be provided about: antibody dilutions, company names, catalogue numbers and clone numbers for monoclonal antibodies; sequences of RNAi and cDNA probes/primers or company names and catalogue numbers if reagents are commercial; cell line names, sources and information on cell line identity and authentication. Animal studies and experiments involving human subjects must be reported in detail, identifying the committees approving the protocols. For studies involving human subjects/samples, a statement must be included confirming that informed consent was obtained. Statistical analyses and information on the reproducibility of experimental results should be provided in a section titled “Statistics and Reproducibility”.

All Nature Cell Biology manuscripts submitted on or after March 21 2016 must include a Data availability statement at the end of the Methods section. For Springer Nature policies on data availability see http://www.nature.com/authors/policies/availability.html; for more information on this particular policy see http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf. The Data availability statement should include:

• Accession codes for primary datasets (generated during the study under consideration and designated as "primary accessions") and secondary datasets (published datasets reanalysed during the study under consideration, designated as "referenced accessions"). For primary accessions data should be made public to coincide with publication of the manuscript. A list of data types for which submission to community-endorsed public repositories is mandated (including sequence, structure, microarray, deep sequencing data) can be found here http://www.nature.com/authors/policies/availability.html#data.

• Unique identifiers (accession codes, DOIs or other unique persistent identifier) and hyperlinks for datasets deposited in an approved repository, but for which data deposition is not mandated (see here for details http://www.nature.com/sdata/data-policies/repositories).

• At a minimum, please include a statement confirming that all relevant data are available from the authors, and/or are included with the manuscript (e.g. as source data or supplementary information), listing which data are included (e.g. by figure panels and data types) and mentioning any restrictions on availability.

• If a dataset has a Digital Object Identifier (DOI) as its unique identifier, we strongly encourage including this in the Reference list and citing the dataset in the Methods.

We recommend that you upload the step-by-step protocols used in this manuscript to the Protocol Exchange. More details can found at www.nature.com/protocolexchange/about.

DISPLAY ITEMS – main display items are limited to 6-8 main figures and/or main tables for Articles, Resources, Technical Reports; and 5 main figures and/or main tables for Letters. For Supplementary Information see below.

FIGURES – Colour figure publication costs $600 for the first, and $300 for each subsequent colour figure. All panels of a multi-panel figure must be logically connected and arranged as they would appear in the final version. Unnecessary figures and figure panels should be avoided (e.g. data presented in small tables could be stated briefly in the text instead).

All imaging data should be accompanied by scale bars, which should be defined in the legend.

Cropped images of gels/blots are acceptable, but need to be accompanied by size markers, and to retain visible background signal within the linear range (i.e. should not be saturated). The boundaries of panels with low background have to be demarked with black lines. Splicing of panels should only be considered if unavoidable, and must be clearly marked on the figure, and noted in the legend with a statement on whether the samples were obtained and processed simultaneously. Quantitative comparisons between samples on different gels/blots are discouraged; if this is unavoidable, it should only be performed for samples derived from the same experiment with gels/blots were processed in parallel, which needs to be stated in the legend.

Figures should be provided at approximately the size that they are to be printed at (single column is 86 mm, double column is 170 mm) and should not exceed an A4 page (8.5 x 11"). Reduction to the scale that will be used on the page is not necessary, but multi-panel figures should be sized so that the whole figure can be reduced by the same amount at the smallest size at which essential details in each panel are visible. In the interest of our colour-blind readers we ask that you avoid using red and green for contrast in figures. Replacing red with magenta and green with turquoise are two possible colour-safe alternatives. Lines with widths of less than 1 point should be avoided. Sans serif typefaces, such as Helvetica (preferred) or Arial should be used. All text that forms part of a figure should be rewritable and removable.

We accept files from the following graphics packages in either PC or Macintosh format:

- For line art, graphs, charts and schematics we prefer Adobe Illustrator (.AI), Encapsulated PostScript (.EPS) or Portable Document Format (.PDF). Files should be saved or exported as such directly from the application in which they were made, to allow us to restyle them according to our journal house style.

- We accept PowerPoint (.PPT) files if they are fully editable. However, please refrain from adding PowerPoint graphical effects to objects, as this results in them outputting poor quality raster art. Text used for PowerPoint figures should be Helvetica (preferred) or Arial.

- We do not recommend using Adobe Photoshop for designing figures, but we can accept Photoshop generated (.PSD or .TIFF) files only if each element included in the figure (text, labels, pictures, graphs, arrows and scale bars) are on separate layers. All text should be editable in ‘type layers’ and line-art such as graphs and other simple schematics should be preserved and embedded within 'vector smart objects’ - not flattened raster/bitmap graphics.

- Some programs can generate Postscript by 'printing to file' (found in the Print dialogue). If using an application not listed above, save the file in PostScript format or email our Art Editor, Allen Beattie for advice (a.beattie@nature.com).

Regardless of format, all figures must be vector graphic compatible files, not supplied in a flattened raster/bitmap graphics format, but should be fully editable, allowing us to highlight/copy/paste all text and move individual parts of the figures (i.e. arrows, lines, x and y axes, graphs, tick marks, scale bars etc.). The only parts of the figure that should be in pixel raster/bitmap format are photographic images or 3D rendered graphics/complex technical illustrations.

All placed images (i.e. a photo incorporated into a figure) should be on a separate layer and independent from any superimposed scale bars or text. Individual photographic images must be a minimum of 300+ DPI (at actual size) or kept constant from the original picture acquisition and not decreased in resolution post image acquisition. All colour artwork should be RGB format.

FIGURE LEGENDS – must not exceed 350 words for each figure to allow fit on a single printed NCB page together with the figure. They must include a brief title for the whole figure, and short descriptions of each panel with definitions of the symbols used, but without detailing methodology.

TABLES – main tables should be provided as individual Word files, together with a brief title and legend. For supplementary tables see below.

SUPPLEMENTARY INFORMATION – Supplementary information is material directly relevant to the conclusion of a paper, but which cannot be included in the printed version in order to keep the manuscript concise and accessible to the general reader. Supplementary information is an integral part of a Nature Cell Biology publication and should be prepared and presented with as much care as the main display item, but it must not include non-essential data or text, which may be removed at the editor's discretion. All supplementary material is fully peer-reviewed and published online as part of the HTML version of the manuscript. Supplementary Figures and Supplementary Notes are appended at the end of the main PDF of the published manuscript.

Supplementary items should relate to a main text figure, wherever possible, and should be mentioned sequentially in the main manuscript, designated as Supplementary Figure, Table, Video, or Note, and numbered continuously (e.g. Supplementary Figure 1, Supplementary Figure 2, Supplementary Table 1, Supplementary Table 2 etc.).

Unprocessed scans of all key data generated through electrophoretic separation techniques need to be presented in a supplementary figure that should be labelled and numbered as the final supplementary figure, and should be mentioned in every relevant figure legend. This figure does not count towards the total number of figures and is the only figure that can be displayed over multiple pages, but should be

provided as a single file, in PDF or TIFF format. Data in this figure can be displayed in a relatively informal style, but size markers and the figures panels corresponding to the presented data must be indicated.

The total number of Supplementary Figures (not including the "unprocessed scans" Supplementary Figure) should not exceed the number of main display items (figures and/or tables (see our Guide to Authors and March 2012 editorial http://www.nature.com/ncb/authors/submit/index.html#suppinfo; http://www.nature.com/ncb/journal/v14/n3/index.html#ed). No restrictions apply to Supplementary Tables or Videos, but we advise authors to be selective in including supplemental data.

Each Supplementary Figure should be provided as a single page and as an individual file in one of our accepted figure formats and should be presented according to our figure guidelines (see above). Supplementary Tables should be provided as individual Excel files. Supplementary Videos should be provided as .avi or .mov files up to 50 MB in size. Supplementary Figures, Tables and Videos much be accompanied by a separate Word document including titles and legends.

GUIDELINES FOR EXPERIMENTAL AND STATISTICAL REPORTING

REPORTING REQUIREMENTS – To improve the quality of methods and statistics reporting in our papers we have recently revised the reporting checklist we introduced in 2013. We are now asking all life sciences authors to complete two items: an Editorial Policy Checklist (found here https://www.nature.com/authors/policies/Policy.pdf) that verifies compliance with all required editorial policies and a reporting summary (found here https://www.nature.com/authors/policies/ReportingSummary.pdf) that collects information on experimental design and reagents. These documents are available to referees to aid the evaluation of the manuscript. Please note that these forms are dynamic 'smart pdfs' and must therefore be downloaded and completed in Adobe Reader. We will then flatten them for ease of use by the reviewers. If you would like to reference the guidance text as you complete the template, please access these flattened versions at http://www.nature.com/authors/policies/availability.html.

STATISTICS – Wherever statistics have been derived the legend needs to provide the n number (i.e. the sample size used to derive statistics) as a precise value (not a range), and define what this value represents. Error bars need to be defined in the legends (e.g. SD, SEM) together with a measure of centre (e.g. mean, median). Box plots need to be defined in terms of minima, maxima, centre, and percentiles. Ranges are more appropriate than standard errors for small data sets. Wherever statistical significance has been derived, precise p values need to be provided and the statistical test used needs to be stated in the legend. Statistics such as error bars must not be derived from n<3. For sample sizes of n<5 please plot the individual data points rather than providing bar graphs. Deriving statistics from technical replicate samples, rather than biological replicates is strongly discouraged. Wherever statistical

significance has been derived, precise p values need to be provided and the statistical test stated in the legend.

Information on how many times each experiment was repeated independently with similar results needs to be provided in the legends and/or Methods for all experiments, and in particular wherever representative experiments are shown.

We strongly recommend the presentation of source data for graphical and statistical analyses as a separate Supplementary Table, and request that source data for all independent repeats are provided when representative experiments of multiple independent repeats, or averages of two independent experiments are presented. This supplementary table should be in Excel format, with data for different figures provided as different sheets within a single Excel file. It should be labelled and numbered as one of the supplementary tables, titled "Statistics Source Data", and mentioned in all relevant figure legends.

--------- Please don't hesitate to contact NCB@nature.com should you have queries about any of the above requirements ---------

Author Rebuttal to Initial comments

We thank the editor for outlining the key referee points to help us with the revision and all reviewers for their insightful and constructive comments to improve the manuscript. We have performed a series of new experiments and revised the manuscript accordingly to address these comments. Below are our responses to the key referee points followed by our point-by-point responses to referees' comments.

Key referee points:

a) use proper substrates that are relevant for in vivo primer formation in the nuclease activity assays, as noted by:

Referee 1:

As far as I can see the nuclease activity of the isolated PPR domain was tested only on a ssRNA substrate, and the assays should therefore be repeated (using the isolated PPR domain and E423P mutant, plus the full-length protein and mutant) on an RNA/DNA hybrid substrate. Understanding the interaction with hybrid may be crucial to understanding the role of the PPR domain in vivo: I would predict that the findings shown in Fig. 3, i.e. the ability of the enzyme to prime DNA synthesis by Poly, are best explained if the PPR domain both degrades ssRNA and stabilizes otherwise transient RNA/DNA hybrid, but does not actually degrade hybrid. In other words creating an RNA 3′ end at a site of hybrid formation. This was not formally demonstrated and is an easy experiment to perform. The final experiments of the paper, showing that the nuclease activity of the PPR domain is required to remove erroneously incorporated bases in mtRNA, raises the intriguing possibility that the PPR domain functions in mismatch surveillance, activating the exonuclease activity when a mismatched base-pair is detected in RNA/DNA hybrid, as the transcription complex advances. As an initial test of this, it would therefore also be worth testing mismatched RNA/DNA

substrates in EMSA for binding and in nuclease assays for trimming activity. Without such experiments, mechanistic understanding of the role of the PPR domain will remain rather minimal.

These are great suggestions. We tested both ribonuclease assays and EMSA using recombinant proteins of the PPR domain, PPR^{E423P}, PolrMT and $PolrMT^{E423P}$ on either a perfectly matched or a mismatched RNA/DNA hybrid. Although the result did not pan out exactly as the reviewer predicted, it did shed light on how PPR or its ribonuclease activity contributes to the primer synthesis and transcription proofreading.

In EMSA, PPR^{E423P} did not show any binding activity (Extended Data Fig. 3c), indicating that Glu^{423} is essential for the PPR domain binding to RNA/DNA hybrid, and explaining why E423P mutant lacks ribonuclease activity. $PolrMT^{E423P}$ still bound to RNA/DNA hybrid, presumably through its polymerase domain (Extended Data Fig. 3d). The PPR domain and PolrMT bound to either matched or mis-matched RNA/DNA hybrids (Extended Data Fig. 3c, d). However, it is difficult to assess their affinity to RNA/DNA hybrids in EMSA because of the cooccurrence of binding and degradation events. U_{29} in RNA/DNA hybrids, particularly in the mismatched one, was rapidly degraded. As a matter of fact, in order to detect binding complexes in EMSA using the mis-matched probe, we had to lower the incubation temperature to 20°C and shorten the incubation time to 15 min.

As expected, neither PPR^{E423P} nor $PolrMT^{E423P}$ showed any ribonuclease activity. The PPR domain and PolrMT degraded ssRNA, RNA in RNA/DNA hybrids or RNA in mismatched hybrids. Degradation profiles of the PPR domain and PolrMT on different substrates can be summarized as below:

1. On a given substrate, these two proteins had similar pattern of degradation, albeit that the PPR domain always showed higher activity than PolrMT. It appears that a PolrMT's PPR domain and its polymerase domain compete for the 3' end of nascent RNA and thereby could mutually inhibit each other.
2. In reactions with either ssRNA or matched hybrid, full-length substrates (5'FAM-RNA, 5'FAM-U_{29} or 5'FAM-U_{29} in U_{29}/dA_{29} hybrid) were only gradually reduced over the course of 30 min period, indicating that the initial formation of substrate-PPR complex, or the removal of first nucleotide from the 3' end is a slow process and potentially the rate limiting step of the degradation.
3. Once the degradation was initiated, the PPR processed on the degradation intermediates from U_{28} to U_{11} in hybrids much faster than these in single-stranded conformation The PPR processed ssRNA (5'FAM-RNA or 5'FAM-U_{29}) slowly from the first nucleotide to the last one. The degradation of U_{29} in an RNA/DNA hybrid by the PPR domain appears biphasic: an initial phase of rapid removal of 18 nucleotides from 3' end, followed by a phase of slow degradation that stopped at the 5^{th} nucleotide to the 5' end, resulting in an accumulation of RNA oligos from 5 nt to 11nt (Fig. 4a).

4. On the mismatched RNA/DNA hybrid, PolrMT proceeded quickly and then completely stopped at the 5th position to the 5'end, resulting a major product of 5 nt oligo (Fig. 4e),
5. The full-length U_{29} in a mismatched hybrid (U_{29}/dC-dA_{28}) disappeared almost immediately in the reaction with the PPR domain, suggesting that the mismatched hybrid is the preferred substrate of the PPR domain for the recognition or binding to initiate the degradation. Consistent with this notion, PolrMT also excised the first nucleotide from U_{29} in the mismatched hybrid much faster than the matched hybrid.

We propose that there is a "tug of war" between the PPR domain and polymerase domain in competing for the 3' terminus of a nascent RNA stand. In the event of a nucleotide misincorporation, the mismatched 3' end can be effectively recognized by the PPR domain and trigger the degradation of RNA strand. The "tug of war" between the polymerase domain and PPR domain may also determine the switch between transcription and primer synthesis: In the event of a lack of factors that stimulate polymerase activity, or the presence of cis-elements on DNA template that hinder the proceeding of the transcription complex, the PPR domain may capture the 3' end of nascent RNA, and initiate the degradation to generate short oligos to prime mtDNA replication.

Referee 3:
2. The exclusive use of single-stranded M13 DNA at the template for all in vitro transcription analyses is a limitation of the study. The in vivo substrate for primer formation is double-stranded mtDNA and, in the case of the leading-strand origin, primers are generated from RNA derived from promoter-specific transcription events. Can the authors show with either the Drosophila or human POLRMT that the exonuclease activity is relevant for primer formation in this more relevant context? This would also allow a better conclusion that the DPPR version of the enzyme has normal transcription activity.

The main finding in our manuscript explains how the ribonuclease activity of PolrMT contributes to the primer synthesis for the lagging strand replication in *Drosophila,* while the mechanism of the leading strand initiation remains unknown. Nonetheless, the PPR domain of human POLRMT also had exoribonuclease activity, which could potentially affect the processivity of the polymerase domain. Therefore, it is an intriguing idea whether the exoribonuclease activity of POLRMT contributes to the leading strand initiation. We thank the reviewer for this insightful suggestion.

We expressed and purified both the full-length human POLRMT and a deletion lacking the PPR domain. We performed *in vitro* transcription assay using a 359 bp fragment spanning the light-stand promoter of human mtDNA as the template. POLRMT generated both the 200 nt run-off transcript and the 120 nt prematurely terminated transcript, which supposedly will be further processed by RNase H1 to prime the heavy strand synthesis. Importantly, the PPR deletion mutant produced the 200 nt run-off transcript only, indicating that the exoribonuclease activity of POLRMT is involved in this premature termination, and

potentially contributes to the leading strand initiation in human mitochondria. We included this data in the revised Extended Fig. 2f-h, and revised the text accordingly on page 11, line 8.

b) provide missing controls in relevant experiments, as noted by:
Referee 1:
b/ Nucleic-acid binding and nuclease assays: the statement in lines 5-7 on p. 6 ("Given that PPR motifs have been proposed to be RNA binding proteins, we hypothesized that PPR might bind to some short DNA sequences to initiate RNA synthesis from these discrete sites on mtDNA") seems fundamentally illogical. If it binds to RNA, why should we expect it to bind also to (ds?)DNA. The experiment of Fig. 2 is nevertheless justified, just not on this basis. In Fig. 2 the authors show data indicating that the PPR domain of PolrMT binds to RNA/DNA hybrid, and may degrade hybrid as well as ssRNA exonucleolytically. This experiment is crucial to the paper, but there are some missing controls. First, the influence of contaminating E. coli proteins has not been rigorously excluded. Data on the purity and integrity of the isolated PPR domain should be more thoroughly presented in Fig. S3 (e.g. using silver-stained SDS-PAGE gels to document the purification steps and final purity). A protein isolate from control cells bearing the expression vector only, treated identically as the experimental isolate, needs to be tested in parallel. Most importantly, all of the EMSA and nuclease assays should be repeated with the isolated PPR domain bearing the E423P 'nuclease-deficient' mutation – as far as I can see only the full-length protein bearing this mutation was tested in the experiment of Fig. 3. This control is especially important since the data of Fig. 2A imply that the PPR domain of PolrMT might have an RNase H-like activity, potentially making it a second RNase H inside mitochondria.

We agree with the referee that the speculation of PPR binding to dsDNA is too much a leap of logic, and hence deleted this sentence in the revision.

Regarding additional controls, we have tested PPR^{E423P}, $PolrMT^{E423P}$ and four other *Drosophila* PPRs proteins, and none of them showed any ribonuclease activity (Extended Fig. 3a, b & 2c). We would like to stress that all recombinant proteins tested in ribonuclease assay were expressed in *E. coli and* purified using the same procedure. *E. coli lysates* were first treated with polyethyleneimine to remove nucleic acids and nucleic acids binding proteins and then subjected to three-step chromatography purification: affinity, anion exchange and size exclusion. In the HPLC-MS analyses, the purified PPR domain had a single peak at 18166.9 da (Extended Fig. 1b, c), the exact molecular weight as predicted, demonstrating its purify and integrity. Nonetheless, per referee's suggestion, we performed silver staining on SDS-PAGE gel, documenting the purify of purified PPR domain (Extended Fig. 1a). Taken together, we believe the influence of contaminating *E. Coli* proteins has been rigorously ruled out.

We would not consider PolrMT a second RNase H, which is classified as a non-specific endoribonuclease. If the PPR domain had RNase H-like activity, we should have detected complexes consisted of the PPR and partially degraded probe in Fig. 2a. Instead, we only detected one species of protein-RNA/DNA complex. All things considered, we have clearly demonstrated that the PPR domain of PolrMT is a 3'-5'

exoribonuclease, which degrades both ssRNA and the RNA strand in an RNA/DNA hybrid (Fig. 2b-e & 4a, b).

Referee 2:
Figure 4 – In order to determine whether the POLMRT restored the steady-state levels of mtDNA and mitochondrial transcripts, wild type control fly strain results should be shown for all data. Otherwise, it is unclear whether the overexpression of transgenic POLMRT is increasing mtDNA and mt-RNA levels above wild type, whereas POLMRTE423P does not.

We thank the reviewer for the suggestion. We have included results of wild type control (w^{1118}) in the revised Fig. 5. We show that PolrMT transgene rescued steady-state levels of both mtDNA and mitochondrial transcripts, as well as *in-organello* mtDNA replication and transcription. On the contrary, PolrMT^{E423P} partially restored the steady-state level of mitochondrial transcripts and *in-organello* transcription, but neither mtDNA level nor *in-organello* mtDNA replication. To avoid redundancy, we do not include images (EdU incorporation assay) of wild type or *PolrMT*^{del} in Fig. 5c, as these are shown in Fig. 1b.

c) strengthen the data to support the PPR domain of POLRMT as the exoribonuclease to synthesize the primer, as noted by Referee 3:
1. From the results in Figure 1D, the authors conclude that the main effect of deleting the PPR domains (lane 4, DPPR) is inability to synthesize shorter RNA transcripts. However, there is very little signal here at all, suggesting an effect on RNA polymerization as well. Thus, it is difficult to conclude that the DPPR enzyme is only deficient in synthesizing short RNA transcripts.

The mutant deleting the whole PPR domain appeared unstable, as its expression level was extremely low in *E. coli* and the purified protein showed little polymerase activity. Thus, we generated PolrMT^{E423P} mutant, which retained the polymerase activity and generated long transcripts only (Fig. 3d). Per reviewer's suggestion, also for a better logic flow, we generated two additional mutants: a deletion removing two PPR motifs (ΔPPRm, a smaller deletion than ΔPPR) and a deletion truncating the whole N terminal region (ΔN). Like ΔPPR, ΔN showed little polymerase activity, suggesting that residues outside of two PPR motifs might be essential for maintaining the structure and activity of the whole PolrMT Enzyme. Of primary significance, ΔPPRm largely retained polymerase activity, but generated long transcripts, resembling T7 RNA polymerase. This result shows that PPR motifs are necessary for synthesizing short RNA transcripts. We replaced the blot of *in vitro* transcription with a new one and included a Coomassie blue stained gel showing all the recombinant proteins used in this assay in the modified Fig. 1e-g, and revised the text accordingly on page 4, line 13.

3. While the data are largely consistent with the PPR domain of POLRMT possessing 3-prime exonuclease activity, the gels demonstrating this directly are somewhat atypical of a processive exonuclease. That is,

there is an expected progressive loss of full-length products over time, but the smaller degradation products do not gradually increase in concert in Figures 2B, 2C and 3A. In other words, the full array of expected degradation products that differ by one nucleotide is observed at the very first timepoint (5 minutes) and do not steadily increase or shift to smaller size ranges over time as expected. How do the authors explain this? Also, a negative control including the reaction mix and the substrate, but not the enzyme, subjected to the longest incubation period in the experiment (30 mins) should be included.

We appreciate the reviewer's critical comments, which have prompted us to further analyze degradation profiles of the PPR and PolrMT proteins. Please refer to our responses to the reviewer 1 on properties of the PPR domain and PolrMT.

We initially used biotin-labeled substrates in the ribonuclease assay and revealed that PPR domain degraded ssRNA from 3'-5'. The reaction mixes were separated on mini-PAGE gels, and then transferred to membrane in order to detect biotin-labeled nucleotides. While the chemiluminescence is more sensitive than fluorescence, the poor retention of small RNA species on membrane caused the lack of smaller degradation products as showed in Fig. 2b. Therefore, we used 6-FAM labeled substrate to directly visualize degradation intermediates in gel, which has greatly improved the detection of short RNA species. We also resolved degradation products on long sequencing PAGE gels to achieve single-nucleotide resolution.

It appears that the initiation of the degradation, or the removal of the first nucleotide from the 3' end is the rate limiting step, as evidenced by the presence of full-length substrates throughout 30 min. The continuous initiation of remaining full-length substrates over the course of degradation reaction explains the continuous presence of longer degradation intermediates in Fig. 2c & d. Nonetheless, there is a slight, but gradual accumulation of degradation intermediates from 5 nt to 10 nt in Fig. 2c, or from single nucleotide to 11 nt in Fig. 2d.

We performed negative controls including the reaction buffer and substrates without any enzyme over 30 min incubation period, and found no degradation. These results were included in the revised Extended Fig. 1d.

Reviewer #1:
Remarks to the Author:
This paper reports the characterization of the PPR domain of the Drosophila mitochondrial RNA polymerase. The domain was found to be required for the synthesis of short, as opposed to long, transcripts on a ssDNA template, conferred binding to RNA/DNA hybrid and had 3′ to 5′ exoribonuclease activity. In an artificial in vitro system, the active exonuclease was needed to enable DNA synthesis, and was also needed for mtDNA replication during oogenesis in vivo. Sequence analysis of mitochondrial RNA indicated

that the exonuclease activity of the PPR domain was needed for proofreading during transcription. Although I have many issues concerning missing controls and data presentation, the main substance of the paper seems to be sound, and its findings are an original and important contribution to the field. In revising the manuscript, the authors should pay attention to the following points. I apologize for their length, but I feel it is important to be specific and detailed as to what needs to be done to get the manuscript into shape, and why.

We thank the Reviewer for thorough and constructive comments.

1. Nomenclature
Although the manuscript is mainly about Drosophila enzymes, the authors use the human nomenclature almost throughout. Flybase is encouraging authors to use official gene and protein names, so I strongly suggest that at first mention of each protein both the mammalian and fly names are given, e.g. mitochondrial RNA polymerase (human – POLRMT, mouse – Polrmt, Drosophila – PolrMT, CG4644) and thereafter use the species-specific nomenclature as appropriate, at each mention.

We used the species-specific nomenclature as the reviewer suggested in the revision.

2. Other definitions
Throughout the manuscript the authors do not distinguish between leading- and lagging-strand priming. Although little is known of either process in Drosophila, they are likely to be fundamentally different, as in almost all other biological contexts, with leading-strand priming almost certainly requiring transcription from a locally unwound dsDNA template and lagging-strand priming taking place on displaced ssDNA either at the fork or elsewhere. Since almost all of the priming-related experiments presented here used a ssDNA substrate, it should be made clearer that the PPR domain of PolrMT is here implicated in lagging-strand priming, with the mechanics of leading-strand priming left open.

We stressed that the PPR domain of PolrMT is required for lagging-strand priming in *Drosophila* in the revision.

3. Experimental findings
a/ Deletion phenotype (Fig 1): larvae with a clean deletion of PolrMT were small (1A), arrested at L3, and showed no puncta of mtDNA synthesis in fat body (1B). The purified protein was active in synthesizing short RNAs on an M13 ssDNA template whereas T7 pol and the ΔPPR variant synthesized mostly long RNAs and the ΔPPR variant had only a very low activity although the latter point is glossed over. Inclusion of the missing ΔNTE-PPR construct in the assay is really needed to address the issue of whether the PPR domain is needed for full polymerase activity as well as to limit the production of long transcripts.

Per reviewer's suggestion, we generated two additional mutants: a deletion removing the whole N terminal region (ΔN) and a deletion removing two PPR motifs (ΔPPRm, a smaller deletion than ΔPPR). Please refer to our responses to the key referee's points above.

b/ Nucleic-acid binding and nuclease assays: the statement in lines 5-7 on p. 6 ("Given that PPR motifs have been proposed to be RNA binding proteins, we hypothesized that PPR might bind to some short DNA sequences to initiate RNA synthesis from these discrete sites on mtDNA") seems fundamentally illogical. If it binds to RNA, why should we expect it to bind also to (ds?)DNA. The experiment of Fig. 2 is nevertheless justified, just not on this basis. In Fig. 2 the authors show data indicating that the PPR domain of PolrMT binds to RNA/DNA hybrid, and may degrade hybrid as well as ssRNA exonucleolytically. This experiment is crucial to the paper, but there are some missing controls. First, the influence of contaminating E. coli proteins has not been rigorously excluded. Data on the purity and integrity of the isolated PPR domain should be more thoroughly presented in Fig. S3 (e.g. using silver-stained SDS-PAGE gels to document the purification steps and final purity). A protein isolate from control cells bearing the expression vector only, treated identically as the experimental isolate, needs to be tested in parallel. Most importantly, all of the EMSA and nuclease assays should be repeated with the isolated PPR domain bearing the E423P 'nuclease-deficient' mutation – as far as I can see only the full-length protein bearing this mutation was tested in the experiment of Fig. 3. This control is especially important since the data of Fig. 2A imply that the PPR domain of PolrMT might have an RNase H-like activity, potentially making it a second RNase H inside mitochondria. As far as I can see the nuclease activity of the isolated PPR domain was tested only on a ssRNA substrate, and the assays should therefore be repeated (using the isolated PPR domain and E423P mutant, plus the full-length protein and mutant) on an RNA/DNA hybrid substrate. Understanding the interaction with hybrid may be crucial to understanding the role of the PPR domain in vivo: I would predict that the findings shown in Fig. 3, i.e. the ability of the enzyme to prime DNA synthesis by Polγ, are best explained if the PPR domain both degrades ssRNA and stabilizes otherwise transient RNA/DNA hybrid, but does not actually degrade hybrid. In other words creating an RNA 3′ end at a site of hybrid formation. This was not formally demonstrated and is an easy experiment to perform. The final experiments of the paper, showing that the nuclease activity of the PPR domain is required to remove erroneously incorporated bases in mtRNA, raises the intriguing possibility that the PPR domain functions in mismatch surveillance, activating the exonuclease activity when a mismatched base-pair is detected in RNA/DNA hybrid, as the transcription complex advances. As an initial test of this, it would therefore also be worth testing mismatched RNA/DNA substrates in EMSA for binding and in nuclease assays for trimming activity. Without such experiments, mechanistic understanding of the role of the PPR domain will remain rather minimal.

To gain more mechanistic understanding on how ribonuclease activity contributes to primer formation and transcription proofreading, we carried out additional experiments per reviewer's suggestions. Please refer to our responses to the key referee's points above.

c/ Regulation of nuclease activity: the exonuclease activity of the PPR domain has to be context-regulated in some way, or else no primer RNA (or any RNA) would ever be made. The nuclease activity should therefore also be measured using a variant of the full-length protein, incorporating a 'polymerase-dead' mutation. The issue of how the exonuclease activity is activated or limited is referred to at the foot of p. 11/top of p. 12, but in a confusing way, since TEFM appears to have no orthologue in Drosophila, so its involvement is irrelevant. Conversely, the experiments presented here relate to the Drosophila enzyme, and the properties of the PPR domain of the mammalian enzyme may differ. I would suggest to make these points clearer. See also next point, 3d.

Currently, no "polymerase-dead" mutations that specifically abolish the polymerase activity have been identified. Given the time frame of this revision, we did not carry out a systemic mutagenesis on PolrMT to identify such mutations. We nonetheless examined the exoribonuclease activity of both PPR domain and the full length PolrMT on RNA/DNA hybrid, a more physiological-relevant context, per reviewer's suggestion. We found that PPR domain had a higher ribonuclease activity than the full length PolrMT, suggesting that there is a "tug of war" between polymerase domain and PPR domain in competing for the 3' end of RNA strand. The outcome of this competition likely determines the switch between continuous transcription and primer synthesis. The PPR domain of human POLRMT also had 3'-5' exoribonuclease activity. We nonetheless agree with the reviewer it is confusing to suggest TEFM as a potential regulator of the PPR, given that it has no homolog in *Drosophila.* We revised the text on page 13, line 4 to clarify this point.

d/ Role in synthesis of short DNA fragments: overexpression in vivo of the wild-type enzyme but not the E423P mutant restored viability and mtDNA maintenance, and supported the production of short mtDNA fragments in organello (Fig. 4C). The authors comment that this shows an intriguing similarity with the synthesis of 7S DNA in mammalian mitochondria. However, since there is no D-loop in Drosophila mtDNA, no previous detection of an abundant short DNA strand, and the molecules synthesized in organello were not in any way mapped, I think this assertion is fanciful and should be dropped. The fact that only short DNA fragments were synthesized suggests that some other component required for processive mtDNA replication is missing or disabled in the in organello system used.

We agree with the referee that the presence of putative 7S-like DNA in *Drosophila* is too speculative, and hence removed this statement in the revision as the reviewer suggested.

e/ Adult phenotype: the experiments on adult phenotype are, indeed, intriguing, although the authors' statement that they were puzzled by the findings is a little misleading, since the subsequent experiments indicate a likely explanation. Maybe it would suffice just to say that the results were 'initially' puzzling. One other small issue concerns the statement that there is little mtDNA replication in the adult musculature. The measurements of copy number reported in Fig. S4 are all indirect: whilst suggesting that the expression of the mutant enzyme has no effect on copy number, it's not very clear why the authors didn't simply measure copy number in isolated thoraxes of males and females of the different genotypes by qPCR, which would have settled the issue more convincingly. Another variable that might have compromised the data of Fig. 5 concerns the expression levels of the inserted transgenes. Western blot data in Fig. S4C probably answers this but needs proper annotation (see point 6q, below). The targeted insertion sites (plus confirmation that these were the (only) insertions detected in the lines) should be indicated somewhere, e.g. in 'Methods', and qRT-PCR from different tissue and stages would help convince me that the mutant phenotype is due to the mutation per se, rather than over or mis-expression.

We revised the text as the reviewer suggested. We also carried out qPCR to quantify mtDNA copy number in flies of different genotypes and included the data in the revised Extended Fig. 4c. Both wild type PolrMT

and E423P mutant transgenes were single copy and inserted into the same landing site using *PhiC31* integrase-mediated transgenesis. We clarified this point in the Method on page 15, line 12. Since overexpression of E423P mutant is toxic, we used inducible Gal4 system to activate the expression of both transgenes in adults only. The overexpression of E423P in ovary led to reduced mtDNA content in the eggs, which supports an essential role of exoribonuclease activity in mtDNA replication. Given that the overexpression of wild type PolrMT did not show any phenotype, we hence ascribed the mutant phenotype to increased transcription error.

5. Methods
a/ Substrates: the authors state that "all 6-FAM and biotin labelled RNA oligos including oligos used as markers were synthesized by IDT DNA" – what is IDT DNA? – I assume it is a company, but please give company name and location in full. The creation of RNA/DNA hybrid substrates for EMSA is not described.

We specified the company name and its location; and included the protocol of preparing RNA/DNA hybrid in the revised Methods on page 22, lines 7-11.

b/ Climbing assay: it isn't clear to me how the final scores were computed. Please be more specific.

We specified how the final score of climbing ability were computed in the revised Methods on page 15, line 30.

c/ S2 cells: The 'Methods' incudes a section on culture conditions for S2 cells – however, as far as I could see, no use was made in the paper of S2 cells.

We deleted the section of S2 cells culture from the Methods.

6. Data quality, figures and statistics
a/ Fig. 1A – scale bars should be shown, and the mt and mutant larval mouthparts should be shown to the same scale in the rh panel.

We included scale bars for Fig. 1a.

b/ Fig. 1B – is anecdotal. Quantitative data should be added: how many animals of each genotype were analysed, how many TFAM puncta were analysed, how many were EdU positive. Are wt and mutant larvae of the same chronological or developmental age, and at what stage (L3 or something else?). Dissection details or a reference should be added to 'Methods'. Why was this experiment not done also on larvae expressing the E423P variant? This is an important experiment to test whether the ribonuclease activity of PolrMT is needed for mtDNA synthesis in vivo.

We included quantifications on TFAM and EdU puncta in Fig. 1b, c and specified the stage of larva used in the experiment in Methods, page 16, line 23. We also replaced the images of EdU incorporation assay with new ones including TFAM staining in wt and E423P expressing larva in the revised Fig. 5c.

c/ Fig. 1C legend – please make clear if/that the numbering refers to the Drosophila enzyme PolrMT. No data are presented on the purity and intactness of the expressed proteins (e.g. by SDS-PAGE), which should be shown in a supplementary figure: standard procedure for validating the data shown in Fig. 1D.

We specified that the numbering refers to *Drosophila* PolrMT. We also included a Coomassie stained gel showing all recombinant proteins used in the in-vitro transcription assay in revised Fig. 1e.

d/ Fig. 1D – once again, only qualitative data from a single experiment are shown. The track for the ΔPPR variant is so faint, whilst that for ΔNTE so black/overexposed, no reliable conclusion can be drawn regarding the role of PPR domain is favouring short RNA products. Some quantitation by phosphorimaging should be added, a longer exposure of the ΔPPR track should be shown, and markers in the size-range of replication primers should be added. Finally, one informative construct is missing, ΔNTE-PPR. and should be included.

Please refer to our responses to the key referee's points and responses to the reviewer's point 3/a on additional deletion mutants. We added the quantification of *in vitro* transcription in Fig. 1g.

e/ Fig. 3D – the legend (and text) do not make clear whether the assay used fly or human Polγ. I can see no sense in combining the fly PolrMT and human Polγ, but the isolation of the fly Polγ is not described anywhere. And was the PPR domain used in this experiment in fact from the human RNA polymerase. Or the fly enzyme? Since the paper is primarily about the fly enzymes, it is extremely confusing to include experiments on the human enzymes without clearly stating that this is the case and justifying the switch of species. The authors should clarify and justify the provenance of the enzymes used here and in each experiment throughout the paper. Given the absence of two relevant accessory proteins in Drosophila (TEFM and Primpol), experiments using the human enzymes are not, in my opinion, really justified here, except as supplementary data to show that the human system behaves similarly (or not) to the Drosophila one (e.g. as in Fig. S3). Note also that Polγ from the two species even has a different subunit composition, and so a chimeric experiment using Polγ from one species and PolrMt (or its PPR domain) from the other is simply not reliable in this context. In short, the main experiment should be repeated using Drosophila Polγ.

To be more physiologically relevant, we tried to use mitochondrial DNA polymerase in the *in vitro* replication assay. We were unable to produce enzymatically active fly POL γ in either *E. coli* or insect cells as reported previously (Bratic, A., *et al. Nat Commun* **6,** 8808, 2015). We hence requested plasmids expressing human POL γ from Dr. Maria Falkenberg, and used human POL γ, the reconstituted heterodimeric complex of POL γ A and POL γ B in the experiment.

The purpose of the *in vitro* replication assay is to test whether small RNA transcripts generated by PolrMT can prime DNA synthesis, not to reconstitute mtDNA replisome *in vitro*. Therefore, it should not matter which kind of DNA polymerase was used in the assay. As a matter of fact, we also used the Klenow fragment in the assay, and it worked.

f/ Fig. 4A – would make more sense if mtRNA levels were normalized to mtDNA. The wt and E423P mutants would then give rather similar transcript levels, while both would be a bit higher than the deletion (though the data might not be statistically significant). Please specify how 'relative level' was normalized: to what? Wild-type?

Thanks for the suggestion. We added the values of mtRNA levels normalized to mtDNA in the revised Fig. 5a. We specified that both mtRNA and mtDNA were normalized to wild type in the revised Figure Legend and Methods.

g/ Fig. 4B – the red signal in the top rh panel (arrowed) is almost invisible. Please supply higher quality images.

We replaced images with new one including TFAM staining and added quantification in the revised Fig. 5c.

h/ Fig. 4C - 'relative level' of what, normalized to what? At least refer the reader back to 'Methods' where this seems to be specified.

We specified that the levels of incorporation were normalized to wild type in the revised Figure Legend and Methods.

i/ Fig. 4A, C – I found the description of the genotypes sloppy and confusingly presented, especially with the inserted commas. Please use standard Flybase symbols, specify the allele of w, clarify whether ac-Gal4 (? Act5C or achaete?) and UAS-PolrMT are really both on chromosome 3 and remove duplication of PolrMT[del] where flies are homozygous. Specify accurately the source of all lines in 'Methods' or in a supplementary table.

Per the reviewer's suggestion, we revised the description of the genotypes in all Figure Legends and included a supplementary table to specify the source of all lines.

j/ Fig. 5B – although it's in 'Methods', please add to the legend that these were male flies

We specified that these were male flies in the revised Figure Legend.

k/ Fig. 5C – please clarify what is meant in the legend by 'time of a misfit'

We clarified this point in the Figure Legend.

l/ Fig. 5D – the pictures are fine as an example, but are of such poor resolution that indicating mt by arrows is pointless. However, the authors should define 'selected mitochondrial area' more precisely in 'Methods'.

We agree with the reviewer that the resolution is not sufficient to accurately outline mitochondrial area. Since the purpose of this experiment is to test whether there are any defects on mtDNA encoded genes, which was directly proved by the data on transcript errors, we decided to remove the SDH-COX staining results for the sake of brevity. This is a moot point now.

m/ Fig. 5 F-H – why are wild-type control flies not shown in these experiments? Please determine and add the values for wt controls.

The wild type PolrMT overexpressing flies behaved identical to wild type control in a series of assays. Therefore, we only compared RNAseq results between PolrMT and PolrMT^{E423P}.

n/ Fig. 5F – this is a busy figure panel which says rater little. I would think it sufficient to appear in the supplementary figures, especially as the functional significance of poly(A) length in mitochondrial transcripts is not properly understood.

We moved this panel to the revised Extended Fig. 5a.

o/ Fig. 5 legend – once again, the genotypes are confusingly presented. Please use Flybase-approved official symbols and add the information on the Gene Switch strain (again using approved nomenclature) to 'Methods'.

We revised the description of the genotypes according to the Flybase and added information on the gene switch line in the Methods.

p/ Fig. S4A – the genotype of the flies is unstated and no controls are shown.

We indicated the genotype of flies in the revised Figure Legend.

q/ Fig. S4C – the 'POLRMT Bac clone' line is not described anywhere else in the paper. Its construction must be described and justified for use in this experiment. Please also specify where proteins were extracted from: L3 larvae, whole adults (male/female?), thoraxes?

This line was published previously (Zhang et. al., 2016 EMBO). We specified the source of this line in the revised Methods. We also specified the tissues and the stages of flies used for the western blot in the revised Methods.

r/ Fig. S5 title – is misleading, since the very next figure shows that it isn't 'unaffected'. I suggest instead 'The ablation of the exoribonuclease activity of PolrMT does not affect the amount of mitochondrial transcription'.

We agree with the reviewer, and revised the title as the reviewer suggested.

7/ References
No major comments, but some rearrangement and rebalancing should be needed, to rectify the problems with the Introduction (see below).

8. Introduction
Although the PPR domain is found in the mitochondrial RNA polymerase from many species, the specific context of Drosophila in the present experiment needs to be more strongly emphasized and justified. Importantly, the replication mechanism of mtDNA in Drosophila is not known in great detail, but appears to differ from the vertebrate one,. This needs to be made clearer and the experiments to be undertaken in the paper should be cast as being aimed at understanding what is happening in the fly model, rather than directly extrapolating seamlessly to the human/vertebrate context. The very first paragraph introduces 'mammalian mtDNA', but instead should first describe animal mtDNA in general, then more specifically Drosophila mtDNA, and finally data that has been obtained in the mammalian context, for comparison. In all cases it should be made clear what is general and what is or may be unique to mammals (or flies). As currently written, the bulk of the section is about mammalian mtDNA without actually saying so, and this is both irritating and misleading to the reader. Note also that evidence for the strand displacement replication model across most of the mitochondrial genome of Drosophila is weak, whilst strand-coupled replication is well documented. The statement about Primpol is nonsensical: since it is found only in chordates, as stated, it makes no sense to explore its prevalence and potential roles in other animal phyla, where it is absent!

Per reviewer's suggestion, we added relevant studies to give a more accurate portrait of the *Drosophila* mtDNA replication mechanism; moved "mammalian mtDNA part" to the Discussion for a better flow; and removed the statement and data on Primpol.

9/ Minor typos

p. 1, line 22 – 'transcriptional', not 'transcripts'
p. 10, line 26 - 'correct', not 'corrected'

We fixed typos.

Reviewer #2:

Remarks to the Author:
Liu et al report on the characterisation of the PPR domain of Drosophila POLMRT as a primase for mtDNA replication and proof-reader for mitochondrial transcription, though its exoribonuclease activity. The authors used a wide range of biochemical experiments to demonstrates these actions and further determined the consequences of disrupting these activities on fly physiology and behaviour. The study is well controlled in most cases and the authors demonstrated awareness of possible confounding factors to their interpretations that were then carefully excluded. Overall, this is a very nice study that is important for the field.

We appreciate the reviewer's commendation on our work and constructive comments that have helped us to improve the manuscript.

Comments and Questions
The authors have found that fly POLMRT acts in a similar manner to mammalian POLMRT. Figure 4 – In order to determine whether the POLMRT restored the steady-state levels of mtDNA and mitochondrial transcripts, wild type control fly strain results should be shown for all data. Otherwise, it is unclear whether the overexpression of transgenic POLMRT is increasing mtDNA and mt-RNA levels above wild type, whereas POLMRTE423P does not.

We added the results of wild type control flies in the revised Fig. 5.

The position of the PPR domain shown in Figure 1C is different to uniport predictions. Could the authors specify the source of domain prediction?

The positions of the PPR domain, and two PPR motifs were annotated based on NCBI conserved domain search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?INPUT_TYPE=live&SEQUENCE=NP_001259853.1). The NCBI prediction is consistent with the predicted 3D structure of PolrMT, which was modeled after human POLRMT (Ringel et.al., *Nature*. 2011, 478:269-73) using SWISS MODEL.

In figure 3C, the intensity of biotin ssRNA level between POLRMTE423P and T7 RNA polymerase are quite similar, however in the line chart next to it, why is the level of biotin-UTP integration of POLRMTE423P much lower than T7 RNA polymerase over the time. Could the authors please explain this difference?

To demonstrate the full range of products generated by PolrMT, we purposely chose a long-exposed image, in which the signal in T7 RNA polymerase lane was oversaturated. To avoid this confusion, we used a properly exposed blot in the revised Fig. 3d.

In figure S5, the authors claimed that ablation of POLRMT's exoribonuclease activity does not affect mitochondrial transcription. However, in figure 4, the expression of POLRMTE423P in POLRMTdel mutant was unable to restore the mitochondrial transcription. Could the authors please explain this opposite result?

In Figure S5, we used an inducible Gal4 system to over-express PolrMT or PolrMT^{E423P} in the wild type, adult flies, to bypass the developmental arrest caused by PolrMT^{E423P} overexpression. Flies of these two genotypes had the same level of mtDNA to begin with, and had a similar level of mitochondrial transcripts after the overexpression, indicating that PolrMT^{E423P} does not affect transcription.

In Figure 4 (now revised Figure 5), PolrMT or PolrMT^{E423P} were expressed in *PolrMT^{del}* mutant flies, which had reduced mtDNA level. Because PolrMT^{E423P} failed to rescue the steady-state level of mtDNA and hence did not fully restore the steady-state level of mitochondrial transcripts in *PolrMT^{del}* mutant background. We included results of normalized mitochondrial transcripts level (to mtDNA level) in the revised Fig. 5a, to further clarify this point.

Minor points
The article has many grammatical errors and would require proofreading.
Figure 1, 4, Figure S3, S4 panel order does not match textual order.
P. 7 line 27 and 29 – There is no Figure S3F
Consider a heading for the discussion section of the manuscript.

We have proofread the manuscript and corrected errors in the text. We also fixed the issues with these figures, and added a heading for the section of discussion.

Reviewer #3:
Remarks to the Author:
The mechanisms underlying the control of animal mtDNA replication remain to be fully elucidated and important to understand not only from a fundamental perspective, but also because defects in this process lead to human diseases. There is a general consensus that mtDNA replication in mammals (studied mostly in humans and mice) is primed by RNA transcripts synthesized by the mtRNA polymerase (POLRMT), but how short primer RNAs are generated in addition to "full-length" genome-sized transcripts has been a controversial area for many years. For the heavy-strand origin a nuclease, RNAse MRP, was proposed to process POLRMT transcripts via an endonuclease activity in the context of a complex RNA-DNA hybrid to generate RNA primers. However, others have since proposed that regulated transcription termination to form a complex RNA/DNA hybrid that is processed by RNAseH1 instead generates the primers. For the light-strand origin, a similar history exists. A primase enzyme was originally proposed, but was not able to be identified, and the Falkenberg group subsequently proposed that POLRMT is also the OL primase, but with no proposed requirement for a nuclease to generate the primers. The current study by Liu et al. is significant because they provide compelling evidence that POLRMT harbors an exonuclease activity that enables it to generate short RNA primers for mtDNA replication. Specifically, they provide evidence that the PPR domains of Drosophila and human POLRMT have intrinsic 3'-to-5' exonuclease activity. In vitro experiments with a truncated form of POLRMT deficient in exonuclease activity (ΔPOLRMT) indicate that this nuclease activity is required for primer formation, but not RNA polymerization activity (i.e. transcription). In flies, in vivo overexpression of a mutant POLRMT lacking only the exonuclease activity

(POLRMTE423P) results in an increase in mtDNA transcription errors and neuromuscular pathology, implying another role of the ribonuclease activity of POLRMT in proofreading to increase transcriptional fidelity. This is an interesting and important study with high relevance to our understanding of mtDNA replication and elucidates two novel functions of POLRMT, nuclease-mediated primer formation and transcriptional fidelity. There are however some concerns that need to be addressed in order bolster the conclusions reached by the authors that are detailed below:

We thank the reviewer for the insightful and constructive comments.

Main points:
1. From the results in Figure 1D, the authors conclude that the main effect of deleting the PPR domains (lane 4, DPPR) is inability to synthesize shorter RNA transcripts. However, there is very little signal here at all, suggesting an effect on RNA polymerization as well. Thus, it is difficult to conclude that the DPPR enzyme is only deficient in synthesizing short RNA transcripts.

Please refer to our responses to the key Referee points above.

2. The exclusive use of single-stranded M13 DNA at the template for all in vitro transcription analyses is a limitation of the study. The in vivo substrate for primer formation is double-stranded mtDNA and, in the case of the leading-strand origin, primers are generated from RNA derived from promoter-specific transcription events. Can the authors show with either the Drosophila or human POLRMT that the exonuclease activity is relevant for primer formation in this more relevant context? This would also allow a better conclusion that the DPPR version of the enzyme has normal transcription activity.

Please refer to our responses to the key Referee points above.

3. While the data are largely consistent with the PPR domain of POLRMT possessing 3-prime exonuclease activity, the gels demonstrating this directly are somewhat atypical of a processive exonuclease. That is, there is an expected progressive loss of full-length products over time, but the smaller degradation products do not gradually increase in concert in Figures 2B, 2C and 3A. In other words, the full array of expected degradation products that differ by one nucleotide is observed at the very first timepoint (5 minutes) and do not steadily increase or shift to smaller size ranges over time as expected. How do the authors explain this? Also, a negative control including the reaction mix and the substrate, but not the enzyme, subjected to the longest incubation period in the experiment (30 mins) should be included.

Please refer to our responses to the key Referee points above.

4. The authors need to do a better job of reconciling and contextualizing their results with prior published results on mtDNA replication priming. For example, others have implicated transcription termination and RNAse H1 processing of a complex RNA/DNA hybrid in priming OH in mammals and no nuclease has been implicated to prime POLRMT-derived transcripts OL. There is also evidence for direct role for mtSSB in priming that should be discussed. In addition, are the results on deleting the NTE (Fig. 1D) consistent with other published findings that have deleted this region of POLRMT (Posse et al. NAR 2014 42:3638)?

Finally, any salient differences between mammalian mtDNA replication priming and Drosophila should be highlighted more. The same should be done for mtRNA processing if relevant. For example, the statement on Page 3, lines 4–5, "all mitochondrial mRNAs and rRNAs are polyadenylated", citing Mercer et al. (2011), is incorrect. According to that study, ND6 transcript is not polyadenylated in humans. Is this true in fly as well? Please clarify.

The main finding of this study explains how PolrMT contributes to primer synthesis for the lagging strand in *Drosophila*. Nonetheless, we carried out the *in vitro* transcription assay as the reviewer suggested in the point #2. We found the PPR domain of human POLRMT, which had exoribonuclease activity, was required for the premature termination near the heavy strand origin. This result suggests that the PPR domain of POLRMT likely contributes to the leading strand initiation. We are not proposing that the PPR domain is involved in further processing the prematurely terminated transcript. Therefore, we don't find our result conflicting with the RNAse H1 model.

We proposed a model of "tug of war", in which the PPR domain and polymerase domain compete for the 3' end of nascent RNA strand and thereby could mutually inhibit each other. Interestingly, a previous paper (Wanrooij et al., 2008) showed that mtSSB inhibits the polymerase activity of POLRMT but stimulates POLRMT-primed DNA replication, which is in line with our model. We thank the reviewer pointing out these studies. We included the relevant literatures and revised the text in Introduction and Discussion to reconcile with previous studies and to contextualize our findings. We also described the differences between mammalian mtDNA replication priming and *Drosophila* in the revised Discussion.

Our result on the NTE domain deletion is consistent with a previous study (*Posse et al. NAR 2014 42:3638)*, showing that the NTE domain of mouse Polrmt is dispensable for transcription initiation *in vitro* and somehow displays higher transcription activity.

All *Drosophila* mitochondrial transcripts have Poly(A) tail. We specified this point in the revision.

5. Wang and Shadel 1999 (PNAS 96:8046) performed N-terminal deletion analysis of yeast mtRNA polymerase (RPO41) and showed it is required for mtDNA stability. Interestingly, they also showed that one deleted version, while fully capable of RNA polymerization, was no longer capable of generating a prominent shorter transcript from an origin-containing template in vivo (see Fig. 3C). Kruszewski and Golik (Biochemistry/Moscow 2016 81:1101) subsequently showed that this deleted region contains two degenerate PPR domains, suggesting that a PPR-associated nuclease activity of POLRMT may be conserved in yeast. These results seem consistent with the current findings and should be discussed.

Thank you for pointing out the oversight. We revised the discussion to include these studies on page 12, line 20.

6. Please provide more detail on the bioinformatic analysis of unprocessed transcripts as this is not a common practice and could be useful to others in the field. For example, specific questions include the following: Which junctions were included? For example, was the Atp8-Atp6 junction included, which encodes parts of both genes and hence likely left uncleaved during mRNA maturation? What was the criteria chosen to call a read "gene spanning"? Is including 1 nucleotide on gene A and 50 nucleotides on a neighboring gene B sufficient to make a call? How was SAMtools configured to count gene-spanning reads instead of reads mapping to intragenic regions? Was any input (a custom list of sequences, filtering criteria, etc.) provided by the authors to SAMtools? Were any parameters modified? Lastly, please provide the number of flies used to isolate mitochondria for RNA sequencing and the number of samples that were sequenced independently (indicate any sample pooling that was used).

Thank you for the comments. We described the bioinformatic analyses and experimental procedure in more details in the revised Methods. We also listed all unprocessed junctions in the Supplementary Table 2. We didn't include the Atp8-Atp6 junction in the analysis. To obtain number of reads spanning adjacent genes, we used featureCounts from Subread (v2.0.0), with the option "-p -s 0 -M -O -R CORE". We used 1 (default) minimum overlap cutoff. A read contains one or more matched nucleotides to two adjacent genes was called as unprocessed transcripts. We would like to caution that different organisms might require different parameters. For example, if the downstream genes start with a "A", the reads including only one additional "A" following the upstream gene should be excluded, because it is more likely resulted from the poly (A) tail. We used 100 adult flies in one group to isolate mitochondrial and total mitochondrial RNA profiles of adult flies were generated in triplicates.

Minor points.
Please indicate the age of larva used in the EdU incorporation assay shown in Figure 1B either in the main text or in the methods section.

We indicated the stage of larva used in the EdU assay in the revised Methods.

In Figure 1B, if possible, it might be better to show another area on POLRMTdel sample that includes at least one nucleus so that EdU incorporation into the nuclear DNA is shown while it is lacking in the mitochondria, i.e., the lack of EdU signal is not due to a total failure of DNA replication (both nuclear and organellar) at the larval stage tested in this lethal POLRMT deletion model.

We replaced the images of EdU incorporation assay, showing EdU incorporation into nuclear DNA in the revised Fig. 1b.

Please indicate the antibody used to stain TFAM.

The TFAM antibody is a custom antibody generated in our lab. We clarified this point in the revised Methods, page 17, line 3.

In the methods, a normal "x" letter should be used, instead of the multiplication symbol "×", to indicate the concentration of solutions. For example, 10x lysis buffer, 2x YT medium, etc. Some arrowheads described in the figure legends are missing in the figures. See figures 3A, S2C, S2D, S3E and corresponding figure legends, for example.

We fixed these issues in the revised Methods and Figure legends.

Since the majority of the results of the study are with Drosophila and the replication function was only demonstrated in flies, should the title be edited as follows (or in some other way) to reflect this? "The PPR domain of mitochondrial RNA polymerase is a ribonuclease required for mtDNA replication in D. melanogaster".

We revised the title to "The PPR domain of mitochondrial RNA polymerase is a ribonuclease required for mtDNA replication in *D. melanogaster*".

Will the mitochondrial RNA sequencing data be made available online?

Yes, all RNA sequencing data generated in this study have been deposited to the Gene Expression Omnibus of NCBI (GEO: GSE 154310, GSE 192549 and GSE 164324).

Decision Letter, first revision:

Subject: Your manuscript, NCB-X45730A
Message:

Our ref: NCB-X45730A

5th February 2022

Dear Dr. Xu,

Thank you for submitting your revised manuscript "The PPR domain of mitochondrial RNA polymerase is an exoribonuclease required for mtDNA replication in D. melanogaster" (NCB-X45730A). It has now been seen by the original referees and their comments are below. The reviewers find that the paper has improved in revision, and therefore we'll be happy in principle to publish it in Nature Cell Biology, pending minor revisions to satisfy the referees' final requests and to comply with our editorial and formatting guidelines.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements in about a week. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Thank you again for your interest in Nature Cell Biology. Please do not hesitate to contact me if you have any questions.

Sincerely,

Jie Wang, PhD
Senior Editor
Nature Cell Biology

Tel: +44 (0) 207 843 4924
email: jie.wang@nature.com

Reviewer #1 (Remarks to the Author):

The manuscript is much improved, and almost all of my comments and criticisms have been satisfactorily addressed. In regard to my first 'key point' earmarked by the editor for attention, the authors conducted extensive further experiments to analyse the activity of the PPR domain and of residue E423 on relevant substrates. Although the results differed from my predictions, they do now support a clear mechanistic hypothesis and I commend the authors on having completed these experiments rigorously and comprehensively. Regarding virtually all other issues the authors have addressed my concerns, made the requested changes, or argued convincingly in their rebuttal letter.

The failure to express functionally active fly Pol-gamma remains a bit puzzling and a minor blemish on the work. The use of the human enzyme is not ideal, but as the authors state in the rebuttal letter, the point was to show that the fly PPR nuclease was needed to produce primers, not to reconstitute a replisome in vitro. They also point out that Klenow worked too (a point that could be included somewhere in the main text or figure legends). The fly polymerase was purified previously by the Kaguni lab which might have been able to supply it, but I find the authors' arguments and data sufficiently sound, if still a little unsatisfying. I think further experimental work isn't needed here.

Two minor textual issues that remain:

1/ Suppl. Table (S4) is now included, but should be mentioned in the main text!

2/ As suggested, the poly(A) length data are now moved to the Extended Figure. But now they are not mentioned in the main text at all. So either they should be mentioned somewhere or else deleted entirely.

My only other comment is that, in a few places, there are minor lapses of the English that will require copy-editing or final language editing by the authors themselves or their associates.

Overall, this is now an excellent paper.

Reviewer #2 (Remarks to the Author):

Liu et al have returned a revised version of their manuscript that addresses all of my main concerns adequately.

There are still some issues with minor points such as the occurrence of grammatical errors/typos (e.g. lines 325, 333 etc), the ordering of panels and text citations (e.g. Fig 1d introduced after 1f), and inconsistent labelling of figures (e.g. interchangeable use of "wt" and "w1118" in Fig 1. a and b).

In addition Figure 7 is a little unclear and could be explained better.

Line 369-70: please provide a reference for the statement: "… as mtDNA is not actively replicated in post-mitotic tissues of adult flies."

Overall, Liu et al have done a very good job at addressing my concerns.

Reviewer #3 (Remarks to the Author):

The authors have done a very thorough job of responding to the reviews, and have addressed all of my concerns. This is an important study that provides significant new insights into the function of mitochondrial RNA polymerase.

Decision letter, final requests:

Subject: NCB: Your manuscript, NCB-X45730A
Message:

Our ref: NCB-X45730A

14th February 2022

Dear Dr. Xu,

Thank you for your patience as we've prepared the guidelines for final submission of your Nature Cell Biology manuscript, "The PPR domain of mitochondrial RNA polymerase is an exoribonuclease required for mtDNA replication in D. melanogaster" (NCB-X45730A). Please carefully follow the step-by-step instructions provided in the attached file, and add a response in each row of the table to indicate the changes that you have made. Ensuring that each point is addressed will help to ensure that your revised manuscript can be swiftly handed over to our production team.

We would like to start working on your revised paper, with all of the requested files and forms, as soon as possible (preferably within one week). Please get in contact with us if you anticipate delays.

When you upload your final materials, please include a point-by-point response to any remaining reviewer comments.

If you have not done so already, please alert us to any related manuscripts from your group that are under consideration or in press at other journals, or are being written up for submission to other journals (see: https://www.nature.com/nature-research/editorial-policies/plagiarism#policy-on-duplicate-publication for details).

In recognition of the time and expertise our reviewers provide to Nature Cell Biology's editorial process, we would like to formally acknowledge their contribution to the external peer review of your manuscript entitled "The PPR domain of mitochondrial RNA polymerase is an exoribonuclease required for mtDNA replication in D. melanogaster". For those reviewers who give their assent, we will be publishing their names alongside the published article.

Nature Cell Biology offers a Transparent Peer Review option for new original research manuscripts submitted after December 1st, 2019. As part of this initiative, we encourage our authors to support increased transparency into the peer review process by agreeing to have the reviewer comments, author rebuttal letters, and editorial decision letters published as a Supplementary item. When you submit your final files please clearly state in your cover letter whether or not you would like to participate in this initiative. Please note that failure to state your preference will result in delays in accepting your manuscript for publication.

Cover suggestions

As you prepare your final files we encourage you to consider whether you have any images or illustrations that may be appropriate for use on the cover of Nature Cell Biology.

Covers should be both aesthetically appealing and scientifically relevant, and should be supplied at the best quality available. Due to the prominence of these images, we do not generally select images featuring faces, children, text, graphs, schematic drawings, or collages on our covers.

We accept TIFF, JPEG, PNG or PSD file formats (a layered PSD file would be ideal), and the image should be at least 300ppi resolution (preferably 600-1200 ppi), in CMYK colour mode.

If your image is selected, we may also use it on the journal website as a banner image, and may need to make artistic alterations to fit our journal style.

Please submit your suggestions, clearly labeled, along with your final files. We'll be in touch if more information is needed.

Nature Cell Biology has now transitioned to a unified Rights Collection system which will allow our Author Services team to quickly and easily collect the rights and permissions required to publish your work. Approximately 10 days after your paper is formally accepted, you will receive an email in providing you with a link to complete the grant of rights. If your paper is eligible for Open Access, our Author Services team will also be in touch regarding any additional information that may be required to arrange payment for your article.

Please note that Nature Cell Biology is a Transformative Journal (TJ). Authors may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. Find out more about Transformative Journals

Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. For submissions from January 2021, if your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route our standard licensing terms will need to be accepted, including our self-archiving policies. Those standard licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

Please note that you will not receive your proofs until the publishing agreement has been received through our system.

For information regarding our different publishing models please see our Transformative Journals page. If you have any questions about costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com.

Please use the following link for uploading these materials:
[REDACTED]

If you have any further questions, please feel free to contact us. Many thanks!

Best regards,

Ziqian Li
Editorial Assistant
Nature Cell Biology

On behalf of

Jie Wang, PhD
Senior Editor
Nature Cell Biology

Tel: +44 (0) 207 843 4924
email: jie.wang@nature.com

Reviewer #1:
Remarks to the Author:
The manuscript is much improved, and almost all of my comments and criticisms have been satisfactorily addressed. In regard to my first 'key point' earmarked by the editor for attention, the authors conducted extensive further experiments to analyse the activity of the PPR domain and of residue E423 on relevant substrates. Although the results differed from my predictions, they do now support a clear mechanistic hypothesis and I commend the authors on having completed these

experiments rigorously and comprehensively. Regarding virtually all other issues the authors have addressed my concerns, made the requested changes, or argued convincingly in their rebuttal letter.

The failure to express functionally active fly Pol-gamma remains a bit puzzling and a minor blemish on the work. The use of the human enzyme is not ideal, but as the authors state in the rebuttal letter, the point was to show that the fly PPR nuclease was needed to produce primers, not to reconstitute a replisome in vitro. They also point out that Klenow worked too (a point that could be included somewhere in the main text or figure legends). The fly polymerase was purified previously by the Kaguni lab which might have been able to supply it, but I find the authors' arguments and data sufficiently sound, if still a little unsatisfying. I think further experimental work isn't needed here.

Two minor textual issues that remain:

1/ Suppl. Table (S4) is now included, but should be mentioned in the main text!
2/ As suggested, the poly(A) length data are now moved to the Extended Figure. But now they are not mentioned in the main text at all. So either they should be mentioned somewhere or else deleted entirely.

My only other comment is that, in a few places, there are minor lapses of the English that will require copy-editing or final language editing by the authors themselves or their associates.

Overall, this is now an excellent paper.

Reviewer #2:
Remarks to the Author:
Liu et al have returned a revised version of their manuscript that addresses all of my main concerns adequately.

There are still some issues with minor points such as the occurrence of grammatical errors/typos (e.g. lines 325, 333 etc), the ordering of panels and text citations (e.g. Fig 1d introduced after 1f), and inconsistent labelling of figures (e.g. interchangeable use of "wt" and "w1118" in Fig 1. a and b).

In addition Figure 7 is a little unclear and could be explained better.

Line 369-70: please provide a reference for the statement: "… as mtDNA is not actively replicated in post-mitotic tissues of adult flies."

Overall, Liu et al have done a very good job at addressing my concerns.

Reviewer #3:
Remarks to the Author:
The authors have done a very thorough job of responding to the reviews, and have addressed all of my concerns. This is an important study that provides significant new insights into the function of mitochondrial RNA polymerase.

Author Rebuttal, first revision:

Reviewer #1 (Remarks to the Author):

The manuscript is much improved, and almost all of my comments and criticisms have been satisfactorily addressed. In regard to my first 'key point' earmarked by the editor for attention, the authors conducted extensive further experiments to analyse the activity of the PPR domain and of residue E423 on relevant substrates. Although the results differed from my predictions, they do now support a clear mechanistic hypothesis and I commend the authors on having completed these experiments rigorously and comprehensively. Regarding virtually all other issues the authors have addressed my concerns, made the requested changes, or argued convincingly in their rebuttal letter.

The failure to express functionally active fly Pol-gamma remains a bit puzzling and a minor blemish on the work. The use of the human enzyme is not ideal, but as the authors state in the rebuttal letter, the point was to show that the fly PPR nuclease was needed to produce primers, not to reconstitute a replisome in vitro. They also point out that Klenow worked too (a point that could be included somewhere in the main text or figure legends). The fly polymerase was purified previously by the Kaguni lab which might have been able to supply it, but I find the authors' arguments and data sufficiently sound, if still a little unsatisfying. I think further experimental work isn't needed here.

We mention that Klenow also worked in the revised Method on page 27, line 20.

Two minor textual issues that remain:

1/ Suppl. Table (S4) is now included, but should be mentioned in the main text!

We reference the Suppl. Table S4, now Suppl. Table 1 in Figure legends and the Method section.

2/ As suggested, the poly(A) length data are now moved to the Extended Figure. But now they are not mentioned in the main text at all. So either they should be mentioned somewhere or else deleted entirely.

We describe the Poly (A) tail result on page 9, line 5.

My only other comment is that, in a few places, there are minor lapses of the English that will require copy-editing or final language editing by the authors themselves or their associates.

We have proofread the manuscript.

Overall, this is now an excellent paper.

Again, thank you for constructive comments and supporting our manuscript.

Reviewer #2 (Remarks to the Author):

Liu et al have returned a revised version of their manuscript that addresses all of my main concerns adequately.

There are still some issues with minor points such as the occurrence of grammatical errors/typos (e.g. lines 325, 333 etc), the ordering of panels and text citations (e.g. Fig 1d introduced after 1f), and inconsistent labelling of figures (e.g. interchangeable use of "wt" and "w1118" in Fig 1. a and b).

We revised the sentence, fixed the typo, and used “wt” in all places.

In addition Figure 7 is a little unclear and could be explained better.

We describe our model in more detail in the revised legend.

Line 369-70: please provide a reference for the statement: ”… as mtDNA is not actively replicated in post-mitotic tissues of adult flies.”

We referenced the Extended Data Fig. 4a, showing that there is little mtDNA replication in adult muscle, one of the post-mitotic tissues.

A previous study (Kauppila et al, 2018 PNAS, Supplemental Fig S1, C&D) demonstrates that there is little, almost none mtDNA turnover in post-mitotic tissues (head and thorax) over 9-days period, after a 4-days pulse of BrdU feeding. The initial labeling is likely due to mtDNA replication in larva stage. However, the authors do not clarify this issue in the paper. Therefore, we decide not to cite this paper.

Overall, Liu et al have done a very good job at addressing my concerns.

We thank the reviewer for constructive comments and supporting our manuscript.

Reviewer #3 (Remarks to the Author):

The authors have done a very thorough job of responding to the reviews, and have addressed all of my concerns. This is an important study that provides significant new insights into the function of mitochondrial RNA polymerase.

We thank the reviewer for constructive comments and supporting our manuscript.

Final Decision Letter:

Subject: Decision on Nature Cell Biology submission NCB-X45730B
Message:

Dear Dr Xu,

I am pleased to inform you that your manuscript, "The PPR domain of mitochondrial RNA polymerase is an exoribonuclease required for mtDNA replication in D. melanogaster", has now been accepted for publication in Nature Cell Biology.

Thank you for sending us the final manuscript files to be processed for print and online production, and for returning the manuscript checklists and other forms. Your manuscript will now be passed to our production team who will be in contact with you if there are any questions with the production quality of supplied figures and text.

Over the next few weeks, your paper will be copyedited to ensure that it conforms to Nature Cell Biology style. Once your paper is typeset, you will receive an email with a link to choose the appropriate publishing options for your paper and our Author Services team will be in touch regarding any additional information that may be required.

After the grant of rights is completed, you will receive a link to your electronic proof via email with a request to make any corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at rjsproduction@springernature.com immediately.

You will not receive your proofs until the publishing agreement has been received through our system.

Due to the importance of these deadlines, we ask that you please let us know now whether you will be difficult to contact over the next month. If this is the case, we ask you provide us with the contact information (email, phone and fax) of someone who will be able to check the proofs on your behalf, and who will be available to address any last-minute problems.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com

Once your paper has been scheduled for online publication, the Nature press office will be in touch to confirm the details. An online order form for reprints of your paper is available at https://www.nature.com/reprints/author-reprints.html. All co-authors, authors' institutions and authors' funding agencies can order reprints using the form appropriate to their geographical region.

Publication is conditional on the manuscript not being published elsewhere and on there being no announcement of this work to any media outlet until the online publication date in Nature Cell Biology.

Please note that Nature Cell Biology is a Transformative Journal (TJ). Authors may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. Find out more about Transformative Journals

Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. If your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route, the journal's standard licensing terms will need to be accepted, including self-archiving policies. Those licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read

the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

If your paper includes color figures, please be aware that in order to help cover some of the additional cost of four-color reproduction, Nature Research charges our authors a fee for the printing of their color figures. Please contact our offices for exact pricing and details.

As soon as your article is published, you will receive an automated email with your shareable link.

If you have not already done so, we strongly recommend that you upload the step-by-step protocols used in this manuscript to the Protocol Exchange (www.nature.com/protocolexchange), an open online resource established by Nature Protocols that allows researchers to share their detailed experimental know-how. All uploaded protocols are made freely available, assigned DOIs for ease of citation and are fully searchable through nature.com. Protocols and the Nature and Nature research journal papers in which they are used can be linked to one another, and this link is clearly and prominently visible in the online versions of both papers. Authors who performed the specific experiments can act as primary authors for the Protocol as they will be best placed to share the methodology details, but the Corresponding Author of the present research paper should be included as one of the authors. By uploading your Protocols to Protocol Exchange, you are enabling researchers to more readily reproduce or adapt the methodology you use, as well as increasing the visibility of your protocols and papers. You can also establish a dedicated page to collect your lab Protocols. Further information can be found at www.nature.com/protocolexchange/about

You can use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

Please feel free to contact us if you have any questions.

With kind regards,

Jie Wang, PhD
Senior Editor
Nature Cell Biology

Tel: +44 (0) 207 843 4924
email: jie.wang@nature.com