EMBO reports

Manuscript EMBOR-2012-36445

In vivo mutagenesis reveals that OriL is essential for mitochondrial DNA replication

Sjoerd Wanrooij, Javier Miralles Fuste, James B. Stewart, Paulina H. Wanrooij, Tore Samuelsson, Nils-Goran Larsson, Claes M. Gustafsson and Maria Flakenberg

Corresponding author: Maria Falkenberg, University of Gothenburg

Review timeline:

Submission date:	27 July 2012
Editorial Decision:	15 August 2012
Revision received:	24 September 2012
Accepted:	26 September 2012

Transaction Report:

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

1st Editorial Decision 15 August 2012

Thank you for the submission of your manuscript to EMBO reports. We have now received the full set of referee reports, which is copied below.

As you will see, all referees acknowledge the beauty of the approach and agree that the findings are interesting and that they merit publication in EMBO reports. None of the referees raises major concerns and no further experimentation is therefore required. However, the referees do raise some minor issues that need to be addressed in the manuscript text, and they ask for additional quantitation and statistical evaluation of the data, which should be included.

Given these comments, we would like to invite you to revise your manuscript, with the understanding that the referee concerns must be fully addressed and their suggestions taken on board. Acceptance of the manuscript will depend on a positive outcome of a second round of review and I should also remind you that it is EMBO reports policy to allow a single round of revision only and that, therefore, acceptance or rejection of the manuscript will depend on the completeness of your responses included in the next, final version of the manuscript.

Revised manuscripts should be submitted within three months of a request for revision; they will otherwise be treated as new submissions. Also, the revised manuscript may not exceed 30,000 characters (including spaces, references, and figure legends) and 5 figures plus 5 supplementary figures, which should directly relate to the corresponding main figure. Shortening of the manuscript

text may be made easier by combining the Results and Discussion section, which may help to eliminate some redundancy that is inevitable when discussing the same experiments twice. Parts of the materials and methods can also be moved to the supplementary information, however, materials and methods essential for the understanding of the experiments described in the main body of the manuscript may not be moved to the supplement and must remain in the main manuscript file. Please do not hesitate to contact me if you have questions regarding manuscript shortening and format. Please also note that p-values, error bars and the number of experiments performed (n) must be defined in the relevant figure legends.

I look forward to seeing a revised version of your manuscript when it is ready.

Yours sincerely,

Editor
EMBO Reports

REFEREE REPORTS:

Referee #1:

Manuscript D-12-00497 by Wanrooij et al presents clear evidence corroborating the importance of OriL in mtDNA replication, providing strong support for the strand-displacement model of replication. First, the authors analyze the frequency of mutations in regions surrounding OriL in mice with mtDNA mutations seeded by an error prone DNA pol gamma to show a decreased incidence of mutations in the stem-loop region that characterizes OriL. This is a clever use of an extensive dataset that reinforces the clear conservation of OriL in human mtDNA. Second, they conduct a rather thorough series of experiments using assays for primer synthesis by mtRNA polymerase at OriL to show that variants in OriL that survive in mice are functional in vitro. Overall, this is an excellent combination of two distinct approaches that should serve to convince any reader that the strand displacement model for mtDNA replication has far greater merit than the bidirectional model proposed by others.

Fig 6 shows the results of a very detailed analysis of mtDNA genomes for potential OriL sequences showing a remarkable conservation. The authors indicate that an analysis of sequences is available on request. It may be best to deposit this in Supplementary material so that this analysis is sure to be preserved. Also, there is a cryptic statement that birds contain an OriL-like sequence near OriH. This is clearly a loaded issue since the Holt lab published a full paper arguing that there is no lagging strand Ori in chicken. It is reasonable that a full analysis of this issue is beyond the scope of the present manuscript, but the authors should raise this issue in the text rather than burying it in the figure.

Minor points:
1. In Fig 2A, lanes 2 and 3 are numbered, but the contents are not identified

2. There are a few typographical errors:
"exluded" in the figure 1 legend
Souhtern on P. 21
The last line of P.11 reads "essential for at OriL"
Page 14 "pyrimdine" is a typo

Referee #2:

Thank you for sending me this manuscript from Maria Falkenberg and colleagues to review. It is an impressive fusion of in vitro and in vivo studies to try and address the question of whether the stem-loop structure referred to as OriL is essential for mammalian mtDNA replication. The area of mtDNA replication is contentious, with no model(s) generally accepted. Unfortunately, as clearly

stated by the authors, we are unable to test the importance of this structure directly by standard recombinant techniques as we are unable to transform mammalian mitochondria. The approach the authors use is laudable and in my opinion highly convincing. Both in vitro and in vivo experiments are performed to an exceptional standard and I have little critical to add. My only comments are extremely minor. First, on p.8, line 12 the authors state '...all except 2 sites were located in the stem region.' I make it 4. Second, there is an interesting naturally occurring T>C polymorphism at position 5752, immediately proximal to the stem:loop in man as shown in Supp Fig1. This is clearly tolerated. It is not in the exact same position as the essential 5751 as illustrated in Fig 5c, but this figure implies that a T is required immediately proximal to the stem-loop which may be a little misleading. In summary, I believe this manuscript shows beyond reasonable doubt that the stem-loop region referred to as OriL is essential for mtDNA replication and is therefore of substantial interest.

Referee #3:

Does this manuscript report a single key finding? Yes.

This paper provides an in-depth view of the existence of OriL as a main initiation point for a second strand synthesis in the strand-displacement model of mtDNA replication.

Is the reported work of significance? Yes

The authors use a three-tiered approach to highlight the importance of OriL to mtDNA replication. First, they take advantage of increase mtDNA mutagenesis in Pol Gamma exonuclease deficient mice to more easily determine a spectrum of mutations within nearly a kb region centered on OriL. Second, they use primase reactions and reconstituted mtDNA replication reactions to determine necessary and sufficient attributes of priming activities of OriL. Finally, the authors use bioinformatics to survey vertebrates for the existence of OriL with a more stringent look based on the criteria for a successful OriL that was derived from the first two approaches.

Is it of general interest to the molecular biology community? Yes

The details of mtDNA replication has been argued and studied for decades and not only is important for the study of mtDNA replication but also lays the foundation for the discovery of key mtDNA replication proteins that could join the growing list of proteins whose genes are mutated in mitochondrial disease.

Is the major finding robustly documented using independent lines of evidence? Yes.

Criticisms:

The paper asks an important question and successfully uses novel techniques to support the authors' hypothesis. There is no doubt that the pending publication of this work will be a prelude to several studies, including the full spectrum of mtDNA mutations in the mutator mouse and a more in depth study of OriL in interesting organisms such as chicken (where OriL has not been properly identified).

In the meantime, the manuscript in its current condition needs modifications to improve the clarity and quality of science without the need for extensive additional experiments. In general, the field would benefit from the authors taking a more quantitative approach to reporting their data and the significance of their differences (see below for details).

1. In the methods and materials (pg. 22), the authors state that 802 clones were sequenced, but in the results section (pg. 8) the authors state that they sequenced 1747 clones. If the number truly is only 802 clones, then there is problem with calling this "saturation" mutagenesis of a nearly 1000 base region. My guess is that the number is really 1747 in which case the reviewer only asks to fix the mistake in the methods and materials.
2. There should be a statistical way to show that the low frequency of mutations that occur at OriL

are not occurring by chance, while showing an assigned p-value.
3. In Figure 2A, using alpha labeled GTP means that there are fewer G's that should be incorporated in the "-6" template as opposed to the "wt" template (3 in the "-6" and 7 in the "wt"). Gamma-labeled ATP (because the first nucleotide would be an A, and the RNA polymerase would retain the triphosphate at the 5' end) would alleviate this unwanted variable within the experiment and allow for a much better direct comparison. Because "-3" and "-4" should incorporate the same number of C's and the signal is different, I would guess that the observation will be valid. Either the authors should state the limitation of the assay or redo the simple experiment.
4. The label of "input template" next to all of the gels is not properly explained. For a paper of this importance, it should be clear how that band is generated. For instance, Figure 2C and 3B indicate that the increase of RNA polymerase increases the intensity of that band; a result whose explanation is not obvious. If the "input template" band is the generation of one circle, then why does RNA polymerase increase the production of that circular product? Given the importance of identifying specific bands such as OriL products, size markers should be visible with the gel.
5. Pg. 11, regarding Figure 3A and similar experiments, there should be quantitation of this data, i.e. the amount of OriL compared to wt normalized for the amount of input template produced. It almost looks like there is more input template produced in lanes c and d than the other lanes, which could change how those lanes are interpreted. Finally, the description that c and d did not have a "significant effect" is not meaningful without proper quantitation. The term "noticeable effect" would not imply statistical characterization but a qualitative assessment, which is more appropriate without further quantitation.
6. Figure 1D- please define the "frequency" in the y-axis.
7. The methods and materials list an EMSA assay for measuring mtSSB binding to ssDNA substrates. I couldn't find the data for that experiment.
8. The protein amounts in the various reactions are very important and should be displayed in the methods and materials so that the reader can easily determine the stoichiometry of DNA and the various proteins. This should not be referred to in supplemental data of a previous paper but rather listed clearly in this report.

1st Revision - authors' response 24 September 2012

We would like to thank the reviewers for careful reading and helpful comments. We have tried to address their concerns to the best of our abilities.

We would like to point out to the reviewer's that the original manuscript was submitted to EMBO journal, but transferred to EMBO reports during the editorial review process. Since the original manuscript was formatted for EMBO Journal, we have now been forced to rearrange the manuscript quite substantially in order to meet the strict space limitations in EMBO reports. For example, we have fused the results and discussion sections in order to save space.

Please find below, a point-per-point list of the different concerns raised.

Reviewer 1.

Specific point 1.
"Fig 6 (figure 5 in the revised manuscript!) shows the results of a very detailed analysis of mtDNA genomes for potential OriL sequences showing a remarkable conservation. The authors indicate that an analysis of sequences is available on request. It may be best to deposit this in Supplementary material so that this analysis is sure to be preserved. Also, there is a cryptic statement that birds contain an OriL-like sequence near OriH. This is clearly a loaded issue since the Holt lab published a full paper arguing that there is no lagging strand Ori in chicken. It is reasonable that a full analysis of this issue is beyond the scope of the present manuscript, but the authors should raise this issue in the text rather than burying it in the figure."

Our reponse:
We have followed the reviewer's suggestion and now include a list of all identified OriL

sequences as supplemental material (Supplementary Table S2).

We have indeed observed a possible OriL sequence in chicken and some information about this is available in supplementary table S2. However, more experimental data are required in order to experimentally verify the activity of this origin *in vivo*. There are some interesting differences compared with other eukaryotes that we need to study and explain, e.g. the initiating nucleotide is a pyrimidine (CTP) and not a purine (ATP/GTP), which is the case in most other vertebrates. We are currently following up the identification of a chicken OriL in a series of experiments, which we expect to perform during the next 6 months, incl. *in vivo* mapping of RNA to DNA transition at chicken OriL and characterization of the chicken POLRMT *in vitro*. We plan to develop these observations into separate study and we have therefore decided not to include these data in the current manuscript. We hope that the reviewer will find this to be a reasonable way forward. In addition, the strict length limitations of EMBO reports make it physically impossible to squeeze in any additional information in the current manuscript!

Minor points:
1. In Fig 2A, lanes 2 and 3 are numbered, but the contents are not identified

Our reponse:
These are empty lanes, separating the molecular weight marker from the experimental data. We now explain this in the figure legend.

2. We have corrected the typographical errors!

Reviewer 2.

Minor point 1.
My only comments are extremely minor. First, on p.8, line 12 the authors state '...all except 2 sites were located in the stem region.' I make it 4.

Our reponse:
We have corrected this mistake in the new version of the manuscript!

Minor point 2.
.., there is an interesting naturally occurring T>C polymorphism at position 5752, immediately proximal to the stem:loop in man as shown in Supp Fig1. This is clearly tolerated. It is not in the exact same position as the essential 5751 as illustrated in Fig 5c, but this figure implies that a T is required immediately proximal to the stem-loop, which may be a little misleading.

Our response:
Thank you for pointing this out! We have modified figure 4C to emphasize the fact that it is the second T from end of the loop (position 5751) that is absolutely essential for initiation of primer synthesis.

Reviewer 3.

Specific point 1.
In the methods and materials (pg. 22), the authors state that 802 clones were sequenced, but in the results section (pg. 8) the authors state that they sequenced 1747 clones. If the number truly is only 802 clones, then there is problem with calling this "saturation" mutagenesis of a nearly 1000 base region. My guess is that the number is really 1747 in which case the reviewer only asks to fix the mistake in the methods and materials.

Our response:
The reviewer is correct and we have fixed this mistake in the methods and materials!

Specific point 2.
There should be a statistical way to show that the low frequency of mutations that occur at OriL are not occurring by chance, while showing an assigned p-value.

Our response:
We have now statistically validated that the frequency of mutations are lower at OriL than in other, surrounding mtDNA sequences. In the supplementary materials and methods section, we now write:
”To test if the observed decrease in mutation load at OriL was statistically significant, we generated 1000 data sets, each by randomly selecting 32 non-OriL sites from our sequencing data. We then compared the mean per-site mutation rate of each simulated data set to that observed for OriL. The observed per-site mutation load in OriL was lower than in 968 of the 1000 random data sets analyzed ($\alpha = 0.05$, 2 tailed t-test), demonstrating that the OriL sequence differed significantly from the surrounding sequences.”

Specific point 3.
In Figure 2A, using alpha labeled GTP means that there are fewer G's that should be incorporated in the "-6" template as opposed to the "wt" template (3 in the "-6" and 7 in the "wt"). Gammalabeled ATP (because the first nucleotide would be an A, and the RNA polymerase would retain the triphosphate at the 5' end) would alleviate this unwanted variable within the experiment and allow for a much better direct comparison.

Our response
As pointed out by the reviewer, fewer radioactive GTPs are incoporated in the -6 and -7 constructs. The number of 32P-GTP that can be incorporated during primer synthesis is reduced from 7 (wt) to 6 (-1), 5 (-2 and -3), 4 nt (-4 and -5), and 3 (-6). In the new version of figure 2 A the relative efficiency of primer synthesis has been calculated, i.e. we have adjusted for the number of 32P-GTP incorporated.

To emphasize this, we now write in the figure legend to fig 2A:
“The relative efficiency of primer synthesis compared to wt was calculated taking into consideration that the number of incorporated [α-32P]-GTPs differs between the individual constructs.”

Please note that in the accompanying fig 2C, we investigate the effects of the same set of mutations on rolling circle DNA replication. In this experiments the strength of the individual lanes can be directly compared and they correlate nicely with the primer assay displayed in fig 2A.

Specific point #4.
The label of "input template" next to all of the gels is not properly explained. For a paper of this importance, it should be clear how that band is generated. For instance, Figure 2C and 3B indicate that the increase of RNA polymerase increases the intensity of that band; a result whose explanation is not obvious. If the "input template" band is the generation of one circle, then why does RNA polymerase increase the production of that circular product? Given the importance of identifying specific bands such as OriL products, size markers should be visible with the gel.

Our response:
We agree with the reviewer that this is an important point to address. To clarify the nature of the replication products, we have included a new supplemental figure (supplemental figure S2). In this figure, we display a rolling circle DNA replication experiments, including a size marker and explain the nature of the individual bands. In the figure legend to this new supplemental figure, we now write:
“Figure S2. Products formed during rolling-circle mtDNA replication in vitro. (A) Schematic illustration explaining the replication products formed by OriL dependent initiation of laggingstrand DNA synthesis. During the first round of DNA synthesis, the OriL-depending laggingstrand products have a length of about 2100 nts, starting from OriL and stopping at the 5′-end of the template strand. This lagging-strand product can be easily observed, since it is clearly distinct in size from the input template. However, as leading-strand DNA synthesis progresses, OriL is exposed a 2nd, 3rd, and maybe even a 4th time. These later lagging-strand initiation events, will lead to full-length DNA products, i.e. these replicatione vents will progress until they reach the 5′-end of the previous lagging-strand replication event at OriL. The fragments will therefore span the entire distance between two OriL sequences on the leading-strand DNA template (about 3900

nts). The length of these lagging-strand DNA synthesis products will thus be nearly identical to the labeled input template and impossible to separate from the input band, which thus appears stronger. (B) Separation of replication products on a denaturing agarose gel (0.7%, w/v). In the absence of POLRMT (lane 2) only leading strand products can be observed. Templates, not used for productive, rolling-circle replication are also labelled in this reaction. This labeling is most likely explained by POLγ idling on the free 3′-end of the input template. In the presence of POLRMT (lane 3), first passage OriL products migrate as a uniqe band of 2100 nts. Later rounds of OriL-dependent synthesis migrate with the same size as the input template. Therefore, the input band appears stronger, in reactions with active OriL-dependent DNA synthesis."
In addition, we have changed all the figures and now clearly state that the OriL band is due to first round initiation at OriL, whereas later rounds of lagging-strand DNA synthesis, migrates with the same size as the template (input).

Specific point #5.
Pg. 11, regarding Figure 3A and similar experiments, there should be quantitation of this data, i.e. the amount of OriL compared to wt normalized for the amount of input template produced. It almost looks like there is more input template produced in lanes c and d than the other lanes, which could change how those lanes are interpreted. Finally, the description that c and d did not have a "significant effect" is not meaningful without proper quantitation. The term "noticeable effect" would not imply statistical characterization but a qualitative assessment, which is more appropriate without further quantitation.

Our response:
Regarding the effect on the input template, please see our response to specific point #4. We have followed the reviewer's suggestion and quantified the replication products using phosphoimaging. We have only quantified the first round of lagging-strand DNA synthesis (the 2100 nt band), since later initiation events will migrate with the same size as the input template.
We have also changed the wording in the manuscript, and now write "no noticeable" effect instead of "no significant".

Specific point #6.
Figure 1D- please define the "frequency" in the y-axis.

Our response:
We have taken away "frequency" and now instead refer to the percent of mtDNA genomes, which should be clearer!

Specific point #7.
The methods and materials list an EMSA assay for measuring mtSSB binding to ssDNA substrates. I couldn't find the data for that experiment.

We have followed the reviewer´s suggestion and removed the indicated text from the Materials and Method's section. The EMSA assay was included in a previous version of the manuscript, but were taken away during the revision of the final draft.

Specific point #8.
The protein amounts in the various reactions are very important and should be displayed in the methods and materials so that the reader can easily determine the stoichiometry of DNA and the various proteins. This should not be referred to in supplemental data of a previous paper but rather listed clearly in this report.

Our response:
We have followed the reviewer's suggestion and now clearly state the protein concentrations in the Methods section (both in the manuscript and in the supplementary material!).

2nd Editorial Decision 26 September 2012

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports. Thank you for your contribution to our journal.

At the end of this email I include important information about how to proceed. Please ensure that you take the time to read the information and complete and return the necessary forms to allow us to publish your manuscript as quickly as possible.

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

If you do NOT want this File to be published, please inform the editorial office within 2 days, if you have not done so already, otherwise the File will be published by default [contact: contact@emboreports.org]. If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

Thank you again for your contribution to EMBO reports and congratulations on a successful publication. Please consider us again in the future for your most exciting work.

Yours sincerely,

Editorial Assistant
EMBO Reports

REFEREE REPORT:

Referee #3:

The authors thoroughly addressed all criticisms from all reviewers and have noticeably improved the quality of the manuscript. I believe that the manuscript is suitable for publication and a significant advancement in the field. I look forward to seeing the final version of the manuscript in press.

