POINT BY POINT RESPONSE

REVIEWER 1

Dear Authors,
Thank you for your work in this area. Overall I found this paper to be a pleasure to read and I have only minor comments listed below.

Response: Thank you for your kind comments and for identifying specific points that will help to improve the quality of our manuscript. Please, find below our responses and changes performed in the manuscript.

Minor Criticisms

Point 1: Please add a statement about ethical approval (e.g. IRB, etc) for this study.

Response: The statement regarding the ethical approval of this project has been added in the text (Experimental section):

"Research samples and their NGS data were dissociated from personal information, the NGS data derived from the clinical samples was anonymized to avoid any personal identification. The project is approved by the IRB at Instituto de Salud Carlos III (PI 47_2020-v2)."

Point 2: The authors mention that the data will be available in SRA/EMBL. Please consider updating this prior to publication.

Response: BAM files of all the samples in this study with the mtDNA reads have been submitted to the EMBL-ENA repository under PRJEB38987 accession number. This accession number has been updated in the manuscript (Experimental section):

"BAM files containing mitochondrial mapped off-target sequencing reads of all samples described in the article is available at ENA EMBL-EBI repository (PRJEB38987)."

Point 3: The authors note that they assumed homoplasmy in these research samples (line 107). Would recommend that they provide a brief comment on that decision.

Response: We assumed homoplasmy in the research samples in order to classify them as m.1555A>G wild type or carrier because they had low mtDNA coverage (Figure 2), which makes heteroplasmy detection unreliable. We have added this brief explanation in the updated manuscript (Experimental section):

"Off-target reads at m.1555 were extracted and samples were classified as wild type or carriers. In research samples we assumed homoplasmy due to their low mtDNA coverage."

Point 4: In the methods (line 130), the authors provide information about primer sequences. Please consider adding more information about how these were designed to exclude NUMTs.

Response: following your suggestion, we have explained in more detail the primer design which maximized the chances of excluding NUMTs (Experimental section):

"We performed a BLASTn (https://blast.ncbi.nlm.nih.gov/) search of the sequence spanning 300 bases downstream and upstream the m.1555 variant to identify highly homologous regions in nuclear chromosomes (e.g. chromosome 5, 7, 9 and 11). For the fluorescence-based genotyping assay, we indicated the best regions to select specific primers, for Sanger sequencing we manually chose primers with mismatches at the 3' end in the homologous nuclear regions".

Point 5: In figure 1, the authors mention library preparation methods as the driver for variation in % off target. Could consider commenting on the different sizes of the libraries (WES vs targeted), or the specific differences in library preparation that might drive this difference.

Response: Thanks for these details. We have rephrased the concerning paragraph including a comment addressing the relationship between the off-target and the library preparation size:

- We point out that in NGS experiments a larger panel design leads to lower off target, which is also observed in our study when comparing a custom gene panel (~135 KB) versus WES (~60 MB).

- Also, we have added a comment about the higher sequencing depth in WES samples, which leads to more mtDNA reads, despite the custom panel samples showed higher % off-target.

- Finally, we have made more explicit how off-target was influenced by controlled experimental conditions. This effect was clear with our custom panel samples in Figure 1C. In the *Experimental Section*, we specify explain what was the main experimental condition (double-size selection) that modified the % off-target in these sets of samples, which enables to optimize the protocol, following vendor's tips.

Please, find below the three highlighted sentences with the changes proposed above (Results):

"The off-target was higher in the custom gene panel samples than in WES (50% versus 15%), which is an expected inverse correlation between this metric and the design-specific size (135 KB and ~60 MB, respectively) in NGS hybrid capture-based target enrichment experiments. A small fraction of the off-target reads were mapped to the mtDNA (0.03% and 0.02%), resulting in 6x and 21x median mtDNA genome coverage, in the custom gene panel and WES samples, respectively (Table 1). The mtDNA coverage is higher in WES samples since they are sequenced at higher sequencing depths. We found a positive correlation between the proportion of off-target reads and the mtDNA coverage (rWES=0.67, p=7x10-21 and rcustomPanel=0.39, p=2x10-13; Fig. 1). In addition to the panel design size, the off-target varied with specific controlled library preparation conditions in each experiment. This variation in off-target per experiment is shown by the four groups of custom gene panel samples (Fig. 1C)."