Reviewer's report

Title: Characterization of APOBEC3 variation in a population of HIV-1 infected individuals in northern South Africa

Version: 0 Date: 02 Aug 2018

Reviewer: Maria Chahrour

Reviewer's report:

In their manuscript titled "Characterization of APOBEC3 variation in a population of HIV-1 infected individuals in northern South Africa", Matume and colleagues report on variants in the APOBEC3 genes in a cohort of HIV-1 infected individuals from South Africa. Identifying variants in APOBEC3 genes and characterizing their population frequencies is important given the biological function of the encoded proteins and the significance of this in terms of HIV-1 pathogenesis. However, the major caveat in the study (that the authors also acknowledge), is the fact that they identify APOBEC3 variants and characterize their frequencies in a population of HIV-1-infected individuals rather than doing so in unaffected individuals. In my opinion this is a major flaw that needs to be addressed. They are comparing allele frequencies in their cohort of affected individuals to several other cohorts, all of which are not infected with HIV-1. The problem is compounded due to their small sample size. For this study to be meaningful, the authors need to identify variants and compare allele frequencies to a cohort of individuals who are not infected with HIV-1 and are from the same population in the Limpopo province. Furthermore, the authors rely on reference haplotype data from the 1000 Genome project which is not ethnically matched to their study cohort (despite some data from African ancestry in the 1000G project). It is more appropriate to use data from the African Genome Variant Project (see Gurdasani et al., 2015 Nature).

I have additional comments that also need to be addressed, as detailed below:

1- The authors should provide more detail on the clinical picture of individuals in their cohort, and also the demographics of the cohort (age, sex, …).

2- In Table 2, in the "Genotypes" column, the authors provide the cDNA variant position but they do not mention that this is the cDNA position. Please include that in the Table description/title, and also include the NM transcript number for each gene since the cDNA positions are in reference to NM transcripts.

3- The first column in Table 2 states "CDS codon position change & rs numbers". Please change this to "Amino acid change and variant ID".
4- In Table 3, in the "Variation" column, the authors indicate amino acid position. Please include that in the title for the column: "Variation (amino acid and its position)" so the readers can easily recognize.

5- In Table 4, first column title needs to be "Amino acid change and variant ID".

6- Same comment for Table 5, first column title.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.