Reviewer’s report

Title: A viral genome wide association study and genotypic resistance testing in patients failing first line antiretroviral therapy in the first large countrywide Ethiopian HIV cohort

Version: 0 Date: 30 Nov 2018

Reviewer: Valeria Micheli

Reviewer's report:

Dear authors,

the issue you addressed is very interesting. The article is well written but I'd like to suggest the following modifications:

Major Revisions

* The baseline viremia values are very high in all groups (TDF, ZDV and d4T), as you can see in table 1 (mean 5.2, SD 0.8 Log10 copies/mL). Viral load at "month 6" of follow-up could be not an appropriate marker of virological failure, especially if it's defined as viral load>150 copies/mL, due to a possible slower decay of viremia for very high viral load values. Please comment this issue in relation to the detection of drug resistance mutations too.

* The authors used too often acronyms that slow down reading the manuscript.

* The authors mentioned the cost-effectiveness of the NFLG assay: please give an estimation of the cost of this test in comparison to population-based Sanger sequencing and better explain the advantages for an extensive application of NFLG in the clinical practice.

Minor Essential Revisions

1. "Material and Methods/Patients". In table 1 the number of GRTs at both month 6 and 12 is lower (47 vs 51, 30 vs 33). What about the lacking GRTs? Assay performance issue?

2. "Material and Methods/Treatment outcome measurements": the first definition of VF is >150 copies/mL and ≤1.000 copies/mL? It's unclear.

3. "Material and Methods/PBSS": maybe 33 failing patients were at month 12 instead of month 2.

4. "Results/Outcome": there are some mistakes in table 2: for TDF n° of patients "died or LTFU"
is 76 instead of 77. For TDF % of patients "died or LTFU+VL>150" is 38.1 instead of 38.2. For TDF % of patients "died or LTFU+VL>150" is 41.3 instead of 41.9 (see also the text). In addition, it's difficult to derive the source of the denominator used to calculate percentages in table 2. Please explain better adding some information (maybe an extra line).

5. Please comment the only statistically significant value of table 2: the higher % of patients treated with d4T presenting fup

6. For table 3: there are some discrepancies between data in table 3 and along the text (for example %: 82.4% instead of 82.6%); p values are different, please clarify. In addition, what does it mean "n=3" in "Others mutations at month 6"?

7. "Acquired DRM detected by PBSS at month 6 and 12": NRTI+NNRTI n=16/20 (80.0% instead 64.0%, as reported in the paper).

8. "Aminoacid changes identified by NFLG": please specify the region involved in insertion mutations (GLIP/GALN/GTLV/GALN): protease or integrase?

9. "Co-receptor tropism": how you did discriminate between CXCR4 tropic virus and CXCR4/CCR5 dual tropic? Maybe did you choose the cut-off FPR <2.0% to identify CXCR4 tropic virus? Are you sure about the presence of two CXCR4/CCR5 dual tropic viruses at baseline, because you mentioned only one before?

Kind regards

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

Yes

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

Yes

**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.
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