Reviewer’s report

Title: Molecular Epidemiology and Hematologic Characterization of δβ-thalassemia and Hereditary Persistence of Fetal Hemoglobin in 125661 Families of Greater Guangzhou Area, the metropolis of Southern China

Version: 0 Date: 16 Dec 2019

Reviewer: Stella Tamana

Reviewer's report:

This manuscript from Jiang F. et al. reports a large-scale epidemiology and hematological study on Hereditary Persistence Fetal Hemoglobin (HPFH) and δβ-thalassemia in Southern China. The research was performed by collection of blood samples from 125,661 couples participating in the pre-gestational thalassemia screening program from which, 654 individuals were found to have elevated fetal hemoglobin (HbF) levels from 11 districts of Guangzhou. The most important finding of this research is that the authors established the prevalence of the Italian nd-HPFH (Aγ-196 C-T mutation) as the most common nd-HPFH in Southern China.

This manuscript reports significant observations about the prevalence and molecular epidemiology of HPFH/δβ-thalassemia in Southern China that will, ultimately, help to improve thalassemia screening programs and genetic counseling in China. Nevertheless, in my opinion, a suitably revised version of this manuscript might be appropriate for publication in BMC Medical Genetics.

Major points

* The Background section is too short and needs to gradually describe in more detail about HPFH and δβ-thalassemia molecular characteristics, phenotype, prevalence and epidemiology in general and in Southern China specifically. Thus, providing the reader the general background of these diseases, why is important to study (i.e. in terms of the disease-specific burdens) and how epidemiology and prevalence studies help to construct guidelines for prenatal screening tests. Furthermore, in my opinion, the first paragraph of the Discussion section should be moved to the Background section. A table listing the most commonly known mutations and their associated phenotypes (please include standardized vocabulary, e.g. HPO codes, if available) of HPFG/δβ-thalassemia in Southern China will greatly enhance the Background section.
* In Results section a summary table listing all the mutations identified in this study along with the number and percentages of people/couples having these mutations and their main clinical symptoms will provide the reader a quick summary of the results, serve as a reference point for prenatal screening tests and provide more information about the genotype-phenotypes associations.

* In line 29-30 of the Results section it says that none of the subjects were from the same family creating confusion on how it is a familial prevalence (line 39-40 of the same paragraph) of Chinese. Please clarify or choose a more suitable word.

* Results section line 44-45: please provide more information on your observation about the significant differences in the prevalence of deletional HPFH/δβ-thalassemia between Guangzhou districts). What is the composition of the population? How the population diversity affects the carrier rate of these disorders?

* I would expect a small discussion (2-3 lines) about the significance of your observation of the homozygous carrier of the Aγ-196 C-T mutation (line 53-54 of the Results section). This is an important finding which you should emphasize its significance.

* Please review the Discussion section and add the respective references when comparing your findings to other studies. As an example, in line 10 of the second paragraph: "Compared to the three types […] than 6% of our study" in comparison to which study?

* Line 32-33 of third paragraph of Discussion: please provide an explanation of why capillary electrophoresis is better than HPLC.

* Line 53 of the fifth paragraph of Discussion: could you suggest possible ways or experiments for exploring the mechanism of γ-thalassemia caused by HBG2-HBG1 gene fusion?

Minor points

* Extra spaces, minor grammatical and spelling errors need to be corrected throughout the paper.

* In Fig.2 legend please add a more detailed caption to the figure.

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.

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