Author’s response to reviews

Title: Associations between male infertility and ancestry in South Americans: A case control study

Authors:

Maria Skowronek (fernandaskowronek@gmail.com)
Tatiana Velazquez (tavalentina@hotmail.com)
Patricia Mut (mut.patricia@gmail.com)
Gonzalo Figueiro (vazfigue@gmail.com)
Monica Sans (mbsans@gmail.com)
Bernardo Bertoni (bbertoni@fmed.edu.uy)
Rossana Sapiro (rsapiro@fmed.edu.uy)

Version: 1 Date: 02 Feb 2017

Author’s response to reviews:

Dr. Matteo Pasini
Editor in Chief
BMC
BMC Medical Genetics
Dear Dr Pasini,

Please find our corrections of the manuscript entitled “Associations between male infertility and ancestry in South American: a case control study” by Skowronek et al. to be considered for publication as a research paper in BMC Medical Genetics. First of all, we would like to thank the reviewers and their suggestions; we believe that they really helped to improve the quality of the manuscript. Secondly we tried to follow the reviewer's indications and deepen the Y chromosome analysis.

In order to show those data we added table 3. We also decided to include supplementary Table 1 with ANOVA analysis. We believe that these changes clarify the results, enhance the way we
showed the data and also reflect the reviewers’ comments. Finally, below you will find the answers to the reviewers trying to follow their suggestions. We thank you for your consideration and look forward to learning of your decision in the near future.

Sincerely,

Dr. Sapiro

Rossana Sapiro M.D. PhD
Histology and Embriology Department
School of Medicine,
Avenida General Flores 2125, CP 11800,
Montevideo-Uruguay,
Tel: (598-2)-9243414/3551,
Fax: (598-2)-9249563,
e-mail: rsapiro@fmed.edu.uy

Reviewer reports:

Reviewer 1: The paper presents an interesting analysis of the influence of both male and female uniparental ancestry on male fertility and sperm parameters. The main novelty of the paper resides in the investigation of these relations in a population with recent admixture history, such a South American population.

The work presented here has a strong rationale and the methods used to address the questions are appropriate. I am however concerned that the Y chromosome genotyping includes a limited number of markers, and specifically haplogroup F(xK) is not further investigated to assign individuals to specific subgroups.

I think that the authors should provide additional typing, mostly of F(xK) subclades, and eventually repeat statistic tests including the sublcades identified.

This will provide a more detailed analysis, without undermining the result of a significant association between F(xK) and sperm morphology.
Other than this, the paper is generally well written, with a clear explanation of the background and aim and a good discussion of the results.

I suggest that the authors address these minor points to improve the readability of their work:

The overall formatting of the paper should be checked, for instance reference numbers are inconsistent through the text, different formats are used for "et al" (sometimes it appears as et al., sometimes italicised), the Y chromosome is sometimes named Y-chromosome and so on.

Answer to reviewer 1

We would like to thank the reviewer’s suggestions, we followed his/her advice regarding subclades and analyzed cases and controls looking for the subclades of Y chromosome haplogroups. Data were added as table 3. Note that in table 1 we added the primers to type the subclades. Figure 1 was modified in order to include subclades. Some references were duplicated so we corrected the issue.

We also changed et al in line 90, 92, 121 and Y-chromosome in lines 167, 173, 215.

Following the reviewer suggestions we made this modifications to the text of the manuscript:

1) line 53: "inherit" modified to "inherited";
2) line 56: I suggest changing "If" with "Whether"; also, "selection" should be "selective";
3) line 69: the authors should change "male factor determination" as it is unclear, maybe specifying if they are referring to the SRY gene, or generally to the importance of the Y chromosome in male fertility and spermatogenesis

We added a sentence and we think it is clearer now.

4) line 71-73: the sentence starting with "The structure of Y chromosome includes a group..." should be clarified

We added a line in order to clarify the importance of the group of the SNPs in chromosome Y (for the purpose of this work)

5) line 100 and below: the authors should acknowledge that these properties of mtDNA apply to the Y chromosome as well, rather than being exclusive of mtDNA

We believe we clarified this issue by adding the sentence: "It is because of its particular properties, (paternally inherited, lack of recombination in most of its length, one-half consists of tandemly repeated DNA and carrying few genes....." before the explanation of the characteristics of mtDNA.
6) line 109: explain or remove "updates: http://www.phylotree.org/[21].";

We decide to remove the word updates: http://www.phylotree.org/[21]

7) line 123: I suggest "Black and Hispanic/Latino men" to be changed with "males of African or Hispanic ancestry"

The suggestion was accepted

8) line 200: "Fisher's exact probability test" should be just referred to as "Fisher's exact test" for the sake of clarity

The word probability was erased

9) lines 211-212: "below normal for WHO" should be clarified, maybe something like "below normal, according to WHO guidelines"

The line was changed, instead we wrote: 38% of them showed values below normal, according to WHO guidelines

10) line 214: although not exactly a gene, I suggest AZFc to be italicized;

We italicized AZFc

11) line 233: I think the word "normal" is needed here: "...more risk (p<0.01) of presenting NORMAL sperm morphology below...";

We agree with the reviewer that we previously accidentally omitted the word normal, it was corrected

12) line 246: "mayor" should be changed to "higher";

It was changed

13) line 261: "probably" should be removed;

The word probably was removed

14) line 270: :"re arrangements" should be "rearrangements";

That was corrected

15) lines 274-275: TSPY should be italicized and lines 284 and 288: USP9Y should be italicised

TSPY and USP9Y were italicized
Reviewer 2: This study examines genetic causes of male infertility by associating genetic information (mitochondrial haplotype, Y-chromosome haplotype) with (1) fertility problems in a case-control comparison, and (2) sperm characteristics among male patients with fertility problems. The first of the two approaches does not detect any differences in haplotype frequencies between infertile men and controls. The second approach yields an apparent association between the Y-chromosome haplotype F(xK) and sperm morphology, however, the test that is implemented (logistic regression) lacks statistical power because it fails to make use of the full information content of the data that appears to be available. While sperm characteristics of patients were measured on a continuous scale (sperm count, motility, morphology), patients were grouped according to whether the fell below or above a certain cut-off value, and then haplotype frequencies were tested as a predictor of this grouping. I do not understand the reasons for choosing such an analysis strategy where most of the information content is discarded before testing. I strongly recommend re-analysing the data with a more powerful test, namely an ANOVA, where haplotype is used as a predictor of the sperm trait that is measured on a continuous scale (sperm counts should probably be ln-transformed to approach normality). This test should reveal with much greater statistical power whether the claimed association is likely to be real or likely to represent a false-positive finding. The latter are very frequent in exploratory studies of genotype-phenotype association studies like the present one. In my opinion, publication of such exploratory studies should not be conditional on reaching statistical significance (since unreplicated positive findings in association studies are predominantly false-positives anyway), but it is essential that statistical tests make full use of the available data.

Furthermore, the study should present more details about the sperm characteristics that were measured. Currently, the sperm traits and their quantification are not described in the methods and the obtained values are hardly visible from the results section and they are not discussed in comparison to reference values from control subjects that do not suffer from fertility problems.

Answer to the reviewer 2:

First of all we would like to acknowledge all your comments and suggestions since they represented a real positive feedback to our job.

Secondly, a) we agree with the reviewer that sperm analysis and terminology was incomplete and may result unclear to scientists that are not familiar with the field so we added a specific paragraph in Material and Methods section. We believe it will clarify this issue.
b) we asked for English correction to an expert and we think the grammar has been improved in this new version of the manuscript.

c) regarding statistic analysis, we have chosen to categorize the data and perform a logistic regression analysis for two reasons; 1) there are cut-off values in semen analysis that in a way (although not completely) predict the men fertility potential. We think they can have more biological significance that the means alone. 2) the values are not normally distributed.

However, in this new version of the manuscript, we decided to performed an ANOVA test (we agree with the reviewer that it may be informative) so we either applied ANOVA plus Tukey or Kruskal-Wallis One Way Analysis of Variance on Ranks when Normality tests failed (D'Agostino & Pearson omnibus normality test). By applying these tests we found that mean±standard error of sperm morphology is lower in Haplogroup F. It is statistically significant when you compare this group against men from R group. We think that this finding support our data that men in group F have more possibilities of present lower percentage of morphological normal sperm in their ejaculates.

Finally these are the modifications we performed following your specific comments:

1. The Abstract would benefit from a substantial revision. I recommend mentioning "mitochondrial haplotype" and "Y-chromosome haplotype" already in line 34.

We have done the changes as requested

Then I would cut the end of the sentence in lines 35-36 (from "looking" to "failure").

We have done the changes as requested

Line 39 should be changed to "compared among haplogroups by ANOVA".

We added by ANOVA in the abstract

Line 41: I suggest "The genotyping confirmed the known admixture…"

We have done the changes as requested

In line 42, "both groups": specify that these are infertile and controls.

We have done the changes as requested

Line 44: specify that this is a Y-chromosome haplotype.

We have done the changes as requested
Line 46: consider replacing "maternal ancestry" with "mitochondrial haplogroups".

We have done the changes as requested

The current "Conclusion" section is mostly redundant and should be replaced with a statement that the possible association between the Y-chromosome haplotype F(xK) and sperm morphology needs further confirmatory testing.

The conclusion was rewritten following reviewer’s suggestions

2. Line 85: I suggest replacing "applied in medicine from the perspective" with "used to study"
We have corrected the issue

3. Lines 116-118: Why is this finding not used to make a specific confirmatory (rather than exploratory) test? Wouldn't this study predict that men of haplotype H have higher sperm motility? This could be tested with the present data.

We tested this possibility with our data but we did not get the same results as Ruiz-Pesini et al. 2000, namely haplogroup mitochondrial H was not underrepresented in men with low amount of motile sperm in ejaculates.

In Ruiz Pesini’s article the cut off the authors used was 50% of motile sperm (referred to WHO guidelines 1992), and we have been used 32% (WHO 2010). We decided to re-analyze the data using 50% and we did not get different distribution of haplogroups between asthenozoospermic men or men with normal motility. We think that these differences may be in part because of the admixture characteristic of southamerican populations. Our population has Amerindian components (haplogroups ABCD) that are absent in the mentioned study.

4. Line 140: should be changed to "where compared among haplogroups"

Done

5. Line 150: I do not understand the logic of this procedure. If all included patients were already tested negative for Y-chromosome microdeletions, why were they tested here again (lines 192-197), and does this double-testing explain why there was only one single positive case (line 214)?

Sorry for the confusion, we did not tested it twice

The Y chromosome microdeletions were tested only as part of the analysis of infertile men when they are azoospermic and a genetic cause is suspected. It is not part of the aim of this work so we decided to deleted from Material and Methods section.

6. Line 159: Provide much greater detail and 7. Lines 208-210: I had great difficulties making sense of these numbers provided. The term "azoospermia" should be defined for the non-
specialist reader (is this sharp zero sperm or <50,000 sperm per ml?). Then, the number provided seem to say that 18% had 0 sperm, 60% ranged between 50,000 and 5 Mio sperm, and 22% ranged between 5 and 13 Mio sperm, with an average concentration of 3.5 Mio among the latter 82%. However, then in Table 5 it looks like 40 patients actually had sperm counts higher than 15 Mio sperm per ml. These results should be presented in a way that allows for no misunderstandings.

A more complete explanation of the semen analysis was provided in the section of M&M as well as the used nomenclature. We think that it helps to understand both, the general description of semen analysis in the 120 infertile men and table 6 and 7 where data are discriminated according to haplogroups.

8. Lines 211-212: 38% and 65% out of how many (120 or 98)?

We explained that it was out of 120

9. Line 212: "below 4% of normal sperm morphology" This is not understandable to the non-specialist reader and no reference to normal values is given.

We changed the phrase and we think that is more clear now after adding an explanation of the references values in M&M section

10. Line 274: explain TSPY

We added the complete name of TSPY

11. Line 282 and 326: Note that, for the non-recombining part of the Y-chromosome, proximity does not predict association like it does for autosomes.

12. Line 288: Note that this should be true for any Y-chromosome association with infertility, because any mutation that leads to unconditional infertility gets immediately removed from the gene-pool.

We added a line explaining that USP9Y is located in a region associated to infertility (the microdeletion AZFa). Similarly with the other articles cited in our manuscript, we found an association with an haplogroup and alterations of the spermogram. This association is not causative, as the reviewer points correctly. We are suggesting that in fact the SNPs that define the haplogroups are also located in the AZFa region and we need to take into account in the future to explore in the Y chromosome.

13. Lines 316-317: unclear. Maybe it is worth noting that in such cases of population admixture, a relatively lower Y-chromosome diversity has been maintained compared to the mitochondrial diversity (where both European and South American genotypes were maintained)

We tried to clarify the paragraph. In Latin America the historical demographic process has similar characteristics in all the countries. The European men mated Native American women
and this process can be traced through the Y chromosome and mitochondrial DNA. In fact, in admixed populations the diversity for these haplogroups is greater than the original populations. The results show the complexity of this phenotype and some genome structure of the F(xK) (Y chromosome and autosomic) could be contributing to develop infertility.

14. Line 618: "see text": no such text is provided anywhere.

Now is explained in M&M section anyway we erased the word see text since after the reviewer’s comments we found them confusing. We left only the reference

15. Tables 5 and 6: Since numbers do not add up to 120 patients, an additional column with "other haplogroups" or "not genotyped" should be added. Note that we added table 3 so now these tables are 6 and 7

We added the columnn of ND (not determined)