Reviewer’s report

Title: Serum Dioxin and DNA Methylation in the Sperm of Operation Ranch Hand Veterans exposed to Agent Orange

Version: 0 Date: 02 Aug 2019

Reviewer: Andrea Baccarelli

Reviewer's report:

This is a generally well conducted, yet small study using a popular genome-scale DNA methylation platform to investigate the associations between exposure to dioxin and sperm DNA methylation in veterans exposed to Agent Orange in Vietnam. Animal models showed that dioxin has reproductive and transgenerational effects. Therefore, investigating sperm DNA methylation is significant. While the study appears well conducted, this is a small study and many if not all of the results presented are not statistically significant after correcting for multiple comparisons. I suggest that the author modify the presentation of their findings and conclusions and/or conduct additional data analysis to determine whether some of the results may indeed be statistically significant as suggested in my comments below.

MAJOR POINTS

Abstract

The abstract is not written as well as the rest of the paper. In particular, it is not immediately clear why TEAD findings were noteworthy. Also, the final paragraph (conclusions) mentions the replication of the Pilsner's study, but that part of the study is not discussed in the results.

Introduction

In animal models, TCDD has been investigated as a transgenerational toxicant, i.e., a chemical that may induce changes in biological mechanisms (including epigenetics) that may be transmitted to future generations. While this is briefly mentioned in the introduction (Lines 100-02), it could be expanded to further motivate the investigation of DNA methylation in sperm.

Line 102: References 12-14 do not appear to be on TCDD. The statement at lines 100-102 should be rephrased or the references changed.

Methods section

Line 124: there is no explanation of what cycles 2, 5, and 6 are. In fact, I do not think there is any explanation that the AFHS was conducted in cycles. Please amend this issue.

Lines 132-135: Based on the text, it appears that the control category (2.21-7.05 pptr) is largely overlapping with the low TCDD category (<8 ppt pptr). It is not clear from the description here why one group is named "control" and the other "low". I tend to suspect that the controls were not involved
in AO spraying/handling operations, but that is not specified in the text. Also, for consistency, it may be helpful to write the upper and lower values of all the four categories instead of using <8 pptr and >20 pptr.

Line 160: would the SVA adjustment also adjust for potential differences in cell type proportions between samples? I understand that sperm mostly includes sperm cells, but finding variable proportions of epithelial and inflammatory cells is not unusual.

There is no mention of non-detectable TCCD values and how they were handled.

It is unclear whether the blood used for TCDD measurements and the sperm samples were collected at the same time. If the two types of collections were spaced apart in time, the TCDD levels may not fully represent the levels concurrent to the sperm collection. In that case, the authors should estimate concurrent TCDD levels using decay/half-life formulas.

Results section
Line 185 - 193: the TEAD data presented now did not convince me that these results are noteworthy. While 4 CpGs are modified in the same gene, that may also be a random occurrence. I strongly suggest performing - in alternative to the EWAS - an analysis using methods to combine adjacent CpGs such as bumphunting, DMRcate or A-clust. Because bumphunting may not work well with sparse CpGs (such as those included in the 450K platform), the latter methods may be preferred. These methods (by adjusting for multiple comparisons) will be able to reveal whether the TEAD results are significant or clustered in the same gene just by chance. Also, if the new analysis confirmed that TEAD were indeed significant, it would be helpful to know how many TEAD CpGs in total are covered on the 450k array.
Lines 205-228: this analysis of 37 CpG candidates from the Pilsner's study must be adjusted for multiple comparisons either using FDR or Bonferroni. Without multiple-comparison adjustment, these results cannot be interpreted as noteworthy.

Suggestions:
-because the authors appear to be interested in candidate analyses, they may want to consider performing an analysis on the aryl-hydrocarbon receptor gene (AhR). TCDD is the strongest AhR ligand and previous studies have shown that AhR RNA expression is downregulated in humans exposed to TCDD. If DNA methylation were to be associated with TCDD exposure, that would be indicative of how AhR expression is regulated/altered by TCDD. Also, the authors may want to consider expanding their analysis to other related genes such as the AhR repressor (AhRR) or other genes on the AhR pathway. Please note that AhR is also strongly activated by components in tobacco smoke, hence this analysis should be carefully adjusted by smoking (including if available measures of duration and intensity of smoking such as pack years).

-there is growing interest in the potential influences of preconceptional/prenatal risk factors on DNA methylation age estimators, such as the Horvath estimators. The authors may want to assess whether TCDD is associated with DNA methylation age in this study.

MINOR COMMENTS
Line 99: "after exposure to" could be changed to "in".
Line 158: I might have missed it, but I believe that "EWAS" has never been spelled out earlier in the text. If so, please do spell it out here.
General comment: I tend to believe that "ppt" is generally used as an abbreviation for parts per trillion. Is there a reason why "pptr" is being used instead?
Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

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 include 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 declare that I have no competing interests.