Author’s response to reviews

Title: Article miRNA expression profiling in normal breast tissues and hereditary breast tumors from BRCA1- and BRCA2-germline mutation carriers in Brazil

Authors:
Danielle Pessôa-Pereira (dpessoap@outlook.com)
Adriane Evangelista (adriane.feijo@gmail.com)
Rhafaela Causin (rhafaela-lima@hotmail.com)
René da Costa Vieira (reneacv@terra.com.br)
Lucas Machado (lucASFariamachado@yahoo.com)
Iara Santana (iarasantana.hcb@gmail.com)
Vinicius Duval da Silva (vinids@gmail.com)
Karen Borba de Souza (kborbasouza@gmail.com)
Renato José de Oliveira-Silva (renatokjso@gmail.com)
Gabriela Fernandes (gabriela.carvalhofernandes@gamil.com)
Rui Reis (ruireis.hcb@gmail.com)
Edenir Palmero (edenirip@yahoo.com.br)
Marcia Marques (mmemsilveira@gmail.com)

Version: 1 Date: 22 Nov 2019

Author’s response to reviews:

Response Letter

1. The section "Statistical analysis" is poorly detailed, the authors must clarify it.
   Response: The statistical analysis of NanoString data was described in “NanoString data analysis” section. However, it was moved to “Statistical analysis” to avoid misunderstandings.

   Indeed, in the section "Results", subsection "miRNA expression profiling of the hereditary breast cancer and normal breast tissues" the authors mentioned an ANOVA test (page 9, line 236) that was not previously referred on section "Statistical analysis" and no data was shown. Moreover, the authors should review the statistical tests performed. The t-test and ANOVA assume a normal distribution of the sample in study, whereas non-parametric tests should be used for non-normal distribution of the sample.
This information was moved to “Statistical analysis”. Regarding the question about normal distribution, according to the white papers of NanoString, the log2-transformed count data usually satisfy the requirements for t-test and ANOVA. Even more dedicated packages, such as NanostringNorm uses t-test and ANOVA. However, it was used the limma package for statistical analysis and the limma code is very robust against non-normality, i.e., it performs a moderated t-statistics and moderated F-statistics (ANOVA). The moderated statistic is the ratio of the expression value to its standard error. This has the same interpretation as an ordinary t-statistic except that the standard errors have been moderated across genes, effectively borrowing information from the ensemble of the targets to aid with inference about each individual microRNA. Finally, the terms ANOVA in the results were replaced by multiple comparisons.

2. In table 1 it would be more informative if the authors include the tumors' histological subtype. Moreover, the authors should revise the figures presented in the table, because they do not match. For example, in BRCA1X group (n=14) the total described on molecular subtype and the total described for TNM evaluation is 15 cases. Regarding SBC group (n=23) the authors present a total of 47 cases on estrogen receptor evaluation.
Response: A new version was did with this information.

3. The authors present a boxplot in figure 3. This type of graphic is not the most informative to demonstrate the comparison between the different groups. Indeed, no statistically difference seem to be depicted between BRCA1X and BRCA2 breast tumors. The authors should perform a more illustrative graph and present the statistical data.
Response: The boxplot was excluded, as suggested. The plot which present the statistical data is qqplot, but it is not much informative in this context. The Heatmaps represent statistical data indirectly, since only the differential expressed microRNAs were included and is more informative because it shows the modulation of these miRNAs.

4. The authors must include in the "Discussion" a paragraph with the major limitations of the study.
Response: This information was included in Discussion lines 372 - 375.