Author's response to reviews

Title: Evaluation of single nucleotide polymorphisms in microRNAs (hsa-miR-196a2 rs11614913 C/T) from Brazilian women with breast cancer

Authors:

JOSÉ JUVENAL LINHARES (juvenallinhares@gmail.com)
MARCOS AZEVEDO JR (marcosazevedojr@gmail.com)
ADALBERTO ABRAÃO SIUFI (fabriciocolacing@iq.com.br)
CRISTINA VALLETA CARVALHO (crisvcarvalho@yahoo.com.br)
Maria Del Carmen Garcia Molina Wolgien (mcarmen.wolgien@gmail.com)
EMMANUELLE COELHO NORONHA (nellycnl@gmail.com)
TATIANA CARVALHO SOUZA BONETTI (tbonetti@unifesp.br)
ISMAEL DALE COTRIM GUERREIRO DA SILVA (ismael.dale@gmail.com)

Version: 2  Date: 1 November 2012

Author's response to reviews: see over
We would like to thank the reviewer and associated editor for the very constructive comments. Our revisions and corrections are outlined below and changes were highlighted in the text attached. This manuscript was reviewed by a professional science editor and by a native English-speaking copy editor to improve readability.

**Answers to Reviewer and Associate Editor**

Title: Evaluation of single nucleotide polymorphisms in microRNAs (hsa-miR-196a2 rs11614913 C/T) from Brazilian women with breast cancer

Referee 1: Zhibin Hu

The authors have addressed some of my initial comments. But I still have some concerns that need to be considered for authors.

1. I am afraid that the small sample size used in this study may lead to a false positive result. Generally, the effect size for a single variant that were suggested by GWAS may less than 1.5. I do not know how did the authors estimate the statistical power?

   The patients in each group were grouped according to their genotypes; patients with one (heterozygotes) or two mutated alleles (homozygotes) were considered as presenting “variant alleles”, and the wild type genotype was considered as the “control allele”. Thus, we calculated the statistical power based on differences in the prevalence of variant alleles (homozygotes or heterozygotes) in the control and study groups, considering an alpha error of 0.05.

2. Meta-analysis by combining the results of this study with other reported studies may improve the reliability of the findings. Although the authors have discussed the results of meta-analysis in the Discussion section, I suggest they include this part as a section of results. Please also present the methodology and results (e.g., forest plot) in detail.

   We think that inserting the meta-analysis in the results might not contribute to the study because they would bring information about a population that is distinct from the one in the present study. There are no studies in the literature including Brazilian or South American women. The methodology and results were revised one more time and described in detail.
3. The power and limitations of this study should be addressed, which may present readers more information on this study.

The power and limitations are well described in the discussion.

Associate Editor:
The paper is significantly improved compared to the previous version. However, there are still issues with this paper:

1. The abstract states: The analyses indicated that individuals carrying the CC genotype in the hsa-miR-196a2 rs11614913 were associated with decreased risk for breast cancer (p: 0.02; OR, 0.61, CI, 0.42-0.89). The T allele (TC/TT versus CC) was associated with increased risk for breast cancer (p: 0.01; OR, 1.52, CI, 1.11-2.08). Why these results are not shown in any table?

We apologize for that mistake. The data presented in the previous version of the abstract were results from a preliminary analysis. That mistake was corrected and we have included the corrected data in the abstract, which is concordant with those presented in the text.

2. It's best to use one genotype group as a reference and calculate ORs for other genotype groups in relation to it. It's incorrect to say that "individuals ... were associated", because it's the CC genotype was associated with ...

We have considered the CC genotype as the reference, and the analysis for the heterozygote and homozygote variants were carried out separately. The results from these analyses are presented in the manuscript.

3. The results should be presented much better. A table should show both allele frequencies and genotype frequencies in cases and controls. Since the set included both European and non-European patients, please present these results for both populations.

We have calculated the allele and genotype frequencies for both Caucasian and non-Caucasian populations in the cases and control groups. The results are included in the manuscript (table 3).
4. Figure 1: The genotype distribution in the CT and BC groups are shown in Table 2; 100% concordance was observed in the results from the samples’ duplicates. The analysis of the frequencies showed a significant difference between the groups (BC and CT) in regards to genotype distribution ($\chi^2: p=0.024$); the mutant homozygous was more frequent in the CT group than in the BC group ($p=0.009$). Why these tables are embedded in the text, how Figure 1 belong here and what exactly it connects to the Figure 1 presented in the end of the paper? What are the p-values and where they come from? These results should be presented in one place - in the table.

The figure showed the clinical stage and histological grades of patients with breast cancer. These data were not included in the table and p-values were not calculated because the control patients are not classified through those characteristics. In accepting the suggestions from the Associated editor, figure 1 was excluded and the data were included in Table 1.

5. Table 2: The HWE deviation in controls shows a problem, likely due to population admixture and the simple $\chi^2$ test can’t be used here. Show the results in your sub-populations separately and prove that there is no HWE deviation in each of the set (Europeans and non-Europeans) or use a logistic regression analysis adjusting for age, ethnicity, etc...

The HWE was calculated for the Caucasian and non-Caucasian subgroups of patients in both groups (control and BC) and the outcomes are included in the manuscript.

6. It’s unclear why TC and TT genotypes were combined, suggesting a dominant effect of T allele (or recessive effect of C allele), while some other studies combined CC and TC genotypes?

We combined the CT and TT alleles because heterozygous and homozygous individuals can usually provide the evaluation of a polymorphic variant precisely because the heterozygotes carry at least one variant allele. To justify and support this combination of genotypes in the analysis, we are attaching two articles on assessment of polymorphism that also used the same type of grouping.
7. There is no reason to designate any of the alleles as "mutants" because it's a common genetic variation without any proven functional significance. The alleles can be called as "variants" or just C or T allele.

We have changed the terminology “mutants” to “variants”, as suggested.