Author's response to reviews

Title: Accuracy of microRNAs as markers for the detection of neck lymph node metastases in patients with head and neck squamous cell carcinoma

Authors:

Ana Carolina de Carvalho (anacarol.oak@gmail.com)
Cristovam Scapulatempo-Neto (cristovamscapula@gmail.com)
Danielle CC Maia (dellymaia@gmail.com)
Adriane F Evangelista (adriane.feijo@gmail.com)
Mariana A Morini (maremorini@hotmail.com)
André L Carvalho (alopescarvalho@uol.com.br)
André L Vettore (andre.vettore@gmail.com)

Version: 2
Date: 15 February 2015

Author's response to reviews: see over
Dear Dr. Sabina Alam,

Thank you for your e-mail concerning our manuscript entitled “Accuracy of microRNAs as markers for the detection of neck lymph node metastases in patients with head and neck squamous cell carcinoma”. The suggestions made by the reviewers have given us the opportunity to review and improve our manuscript.

As requested in the decision letter, we are resubmitting a revised version of the manuscript. We hope to have been able to address all the points pointed by the reviewers. All the reviewers’ suggestions were incorporated in the revised version of the manuscript and they are highlighted in the text. Below you can find a point-by-point response to all the reviewers’ concerns. Adriane Feijó Evangelista was included as an author of this manuscript because she was responsible for some statistical analyses done to answer reviewers’ questions, which were incorporated in the revised manuscript.

I believe this manuscript meets the stringent requirements of high scientific quality and significance, originality, and priority to be acceptable for publication in BMC MEDICINE. Further, since all the reviewers’ queries could be answered, I would like to ask you to consider this manuscript for publication in BMC MEDICINE.

Sincerely,

Andre L. Vettore

Federal University of São Paulo

São Paulo, Brazil
Responses to reviewers’ comments:

Reviewer 1:

1. This study analyzed the miRNA expression in metastatic lymph node and non-metastatic lymph node of HNSCC. Authors claimed that miR-203 and miR-205 were potential markers for detecting metastatic nodes in the prospective sample cohort retrieving from fine needle biopsy. This study is quite interesting. As the candidate miRNAs including miR-200 family, miR-203 and miR-205 can target ZEB and Snail to repress stemness and EMT, these miRNAs are supposed to be down-regulated during tumor progression. The concept that a repressor for tumor progression can be validated as a metastatic marker appears opaque. As the miRNAs identified from the comparison of this work would stand only for epithelial specificity, this reviewer queries why that a panel of highly expressed oncogenic miRNAs in HNSCC can’t be detected by current approach. These serious concerns need to be addressed by comparing the expression in primary tumor tissues from metastasis cases and non-metastatic primary tumor. The clues may mechanistically link the expression profile of markers to invasion, colonization and metastasis of HNSCC.

Since the aim of the present study was to compare the expression level of microRNAs between metastatic and non-metastatic lymph nodes, which basically means lymph nodes with and without epithelial deposits, we were comfortable in accepting and expecting that the mostly differentially expressed microRNAs between both sample situations were epithelial markers. Our goal was not to find a molecular marker that would signalize that a patient was more likely to harbor metastases, but to identify microRNAs as markers for the detection of metastatic deposits in the lymph nodes. So, this is why we only compare lymph nodes harboring tumor cells vs. non-metastatic ones (according to IHC). We agree with the reviewer that if we had compared primary tumor samples from patients with or without metastases, we should expect to find oncogenic miRNAs differentially expressed between both groups, but this was not the objective of this study. We focused our analysis in lymph node samples and not primary tumor specimens.
We do not agree with the affirmation that “The concept that a repressor for tumor progression can be validated as a metastatic marker appears opaque”. As discussed in the manuscript, based on findings of other studies, we can speculate that the presence of miR-200 family members (miR-200b-200a-429 and miR-200c-141) in the epithelial tumors should contribute for the downregulation of ZEB1 and ZEB2, preventing the inhibition of E-cadherin, avoiding cell differentiation and hindering the promotion of cell migration and invasion via EMT. In other words, these microRNAs should ensure the maintenance of the epithelial phenotype. In sharp contrast, in invasive tumors, it is expected to find a downregulation of these markers in the tumor invasion front, since cells in this region should lose epithelial characteristics and acquire mesenchymal phenotype, enabling them to invade adjacent tissues or vases. However, when reaching the lymph nodes, tumor cells need to re-express epithelial factors to allow the colonization of the new site. Thus, these microRNAs should return to the upregulated state. In conclusion, in the primary tumor tissues, the presence of these markers is not associated with the metastasis status because in these tissues they are contributing for the maintenance of the epithelial phenotype, being expressed in all epithelial cells. For this reason, the expression levels of these microRNAs in primary tumors are not useful as markers for metastasis. On the other hand, the detection of expression of these markers in the lymph nodes reflects the presence of epithelial cells in this lymphoid structure, indicating the presence of metastatic cells. So, in the lymph nodes, the presence of these microRNAs could be used as a marker for the presence of metastases.

2. Table 1: Why FDR-adjusted p values for all miRNAs are 0.1259.

FDR, when enabled in this kind of analysis, controls the expected proportion of falsely rejected hypotheses by removing outliers and reducing the chance of incorrect results (false positives) (Benjamini & Hochbergh, 1994). This analysis allows the association of each discovery to a statistical confidence measure. The procedure consists of sorting the p-values in ascending order, and then dividing each observed p-value by its percentile rank to
get an estimated FDR. In this way, small p-values that appear far down the sorted list will result in small FDR estimates, and vice-versa (Noble, 2009). Because it is a confidence value, the p-adjusted values can be the same for multiple genes, and this is more likely to occur for samples of the same experimental group. Results from other studies also presented similar values of “FDR-adjusted p-values” for groups of similar “raw p-values” (Hui et al. 2010, Cuk et al., 2013).

Important to mention that, as discussed below (reviewer 2, number 4), the t-test used to evaluate the significance of the findings in the discovery set is a parametric test and it is not recommended for evaluation of not normally distributed data set. For this reason, the statistical analysis was revisited, and we performed a non-parametric test (Rank Products) was applied and new selection criteria were adopted (fold change >2 and p-value < 0.05). The FDR-adjusted p-values were not included in the new criteria adopted. Noteworthy, even with the new statistical analysis, the main findings of the manuscript were confirmed.

3. Methods

a. Why “mean Ct of U6 & U47” was used as internal control? Is there any literature support for this sort of mixing.

Based on the proposal of Vandesompele et al. (2002), which advocated that the average expression of two or more reference genes should be used to normalize the mRNA expression across different samples, we decided to use the same approach to determine the microRNA expression level in the samples included in this study.

In the last years, several studies have been conducted using the same approach. For example, in Head and neck tumors, Barker et al., (2009) used the expression average of RNU6B, RNU48 and RNU44 to normalize their expression data, while the choice of Ganci et al., (2014) was the expression average of RNU44 and RNU48.
We used the DataAssist software (Life Technologies) to analyze the data obtained with the TaqMan Human MicroRNA Arrays and selected as internal controls the microRNAs presenting the most stable expression in the two sample groups evaluated (positive and negative lymph nodes from HNSCC patients). According to this analysis, the most stably expressed controls were U6 and U47. For these reason, the expression average of these two microRNAs were used to normalize the expression data of all microRNAs evaluated in this study.


b. What is the meaning of “the mean Ct of 10 non-metastatic lymph nodes was used”. Is there any criteria? Are these nodes randomly selected?

The $2^{-\Delta \text{Ct}}$ method used to calculate fold-change values requires a control group to be used as reference. Thus, the fold-change expression of a given “marker” in case samples will always be compared to the expression of the same “marker” in the reference group (relative expression). Therefore, we randomly selected 10 samples from patients whose all lymph nodes resected were negative for metastases (according to the step-sections and IHC for cytokeratins) and used them as reference group in the $2^{-\Delta \text{Ct}}$ analyses.

This statement was better clarified in the revised version of the manuscript.
4. How come that in retrospective cohort more than 50% of 48 T1-T3 tumors were found to have neck metastasis after surgery? Is there any error in clinical diagnosis?

There was no error in the clinical diagnosis.

The main objective of this study was to analyze lymph node samples and identify microRNAs able to discriminate metastatic and non-metastatic nodes. So, the selection criteria of cases included in this study was directed, as much as possible intentionally design to select 50% of the cases with positive lymph node cases (harboring metastatic cells) and 50% of negative lymph nodes. Our final numbers were 47.9% cases with negative lymph nodes and 52.1% of cases with positive lymph nodes.

5. Fig. 2D is not an intra-nodal involvement. The background tissue seems to be extracapsular.

As seen in the image below (100X magnification), this lymph node containing isolated tumor cells presents some signs of atrophy. However, the metastatic deposit is within the lymph node capsule, thus, although atrophic, the background tissue is lymphoid.
Immunohistochemistry staining for cytokeratins (M3515, clone AE1/AE3, Dako) in a lymph node containing isolated tumor cells (100X magnification).

6. Serious concerns for Fig. 3:

a. Text described M > NM for more than 2-fold, the illustrations of miR-513, miR-484, and others do not look so. Contrasting analysis should be done first to get rid of noise and enrich the important data. The color range scale should also be shown.

As mentioned above, the discovery set data was re-analyzed. Sixty-one presented a p<0.05 (Rank Products) and, of these, 47 showed at least 2-fold upregulation in all four metastatic samples in comparison to the non-metastatic ones were selected. Thereafter, a new heatmap was constructed using a non-supervised hierarchical clustering analysis with the
average ΔCt values of these 47 microRNAs in the 4 metastatic and 2 non-metastatic lymph nodes.

As suggested, a color range scale was added in the Figure 3. The color pattern was changed, which allowed a reduction of the noise and a better visualization of the relevant data. We believe that this new heatmap version shows more clearly the differences in the expression level of the 47 selected microRNAs between non-metastatic and metastatic lymph nodes.
Figure 3: Heatmap representations of the 47 differentially expressed microRNAs with fold-change ≥ 2 and p-value < 0.05 (Rank Products) in the comparison between 4 metastatic lymph nodes (M: 7A, 1A, 8A and 9A) and 2 non-metastatic lymph nodes (NM: 4C and 5C).
resected from patients with T2N0 tongue squamous cell carcinoma. Non-supervised hierarchical clustering plotted based on the average ΔCT values. Up-regulated and down-regulated miRNAs are shown as red and blue, respectively. The columns represent samples and the microRNAs differentially expressed are shown in the lines.

b. As no important oncogenic miRNA was found by current strategy, other algorithm should also be considered.

Since the main objective of the study was to analyze lymph node samples and to identify microRNAs able to discriminate metastatic and non-metastatic nodes, we focused our analysis in identifying microRNAs with high fold-change levels in positive lymph nodes in comparison to non-metastatic ones. For this reason, we adopted a very stringent criteria to select the microRNAs that should be submitted to the validation step: only microRNAs with p-value <0.05 and more than 100-fold increment in the expression level in all 4 metastatic lymph nodes were considered. Noteworthy, the criteria adopted did not take into account the function of the microRNA, but only the high fold-change in the expression level.

We agree that other relevant microRNAs, including oncogenic microRNAs, could be left out of the validation steps because they did not reach the stringent criteria (100-fold and p-value <0.05 in all 4 metastatic samples). For instance, oncomiRs such as miR-147 and miR-183 are present in the list of microRNAs with at least 2-fold upregulation and p-value <0.0.5 in the comparison between metastatic and non-metastatic lymph nodes (discovery set). However, these microRNAs had fold-change levels of 6-fold and 80-fold, respectively. So, due to this “low” fold-change increment, this oncogenic microRNAs were not selected for the validation step.
c. Since the expression profile between NM samples is still variable, the sample size for NM should be expanded.

It is true that the expression data from the NM samples in the PCR array assay present some variation, but it is important to take in account that, this PCR array analysis was used as a screening step (discovery set) for the selection of microRNAs to be evaluated in a larger cohort (validation set). We agree that this screening method adopted was not perfect but it is efficient enough to allow us to pick up some candidates that could be validated in the validation set. Thus, the small fluctuations observed in the NM cases do not avoid the selection of good candidates and we do not think the expansion of the sample size for NM in the discovery set will have an impact on the main findings of the study.

d. Why miR-205 in Table 1 is in borderline significance, while the AUC is 1 in Fig.

It was a typo in Table 1. The correct p-value of miR-205 was 0.0032. However, due to the re-analysis of the discovery dataset using the non-parametric statistical test, all p-values have changed as seen in the new version of Table 1.

7. The M sample number in discovery group should also be expanded.

As argued before, the discovery group was evaluated by PCR array in a screening step for the selection of microRNAs to be evaluated in a larger cohort (validation set). We believe the screening method adopted was efficient enough to allow the selection of candidates that were validated in the validation set. Here again, we do not believe the expansion of the sample size for M in the discovery set will have an impact on the main findings of the study.
8. Why in Fig. 5 only 13 NM samples were analyzed? Ten samples have been used for standardization? The writing of this manuscript should be more precise and clear, particularly in Materials and Methods.

The total number of negative lymph node samples included in the study was 23. Of these 23, the average expression of 10 cases was used as reference group in the \(2^{-\Delta \Delta CT}\) calculations. The expression level of the remaining 13 negative lymph nodes was compared individually with the reference group (10 negative lymph nodes) to enable the calculation of specificity levels.

This was clarified in the revised version of the manuscript.
Reviewer 2:

1. Comments: The article is well written and it is important to the field of HNSCC. There is work done by number of groups highlighting the importance of miRNA to diagnose HNSSC. Please include the following recent publications that cover the importance of miRNA regulation in HNSSC. Please reference the following publications.
   These references were included in the revised version of the manuscript.

2. HNSCC is 5 different subtypes of cancers and as such it is very heterogeneous. Both biologically and clinically these cancers differ based on the anatomical site of origin. Increasing there is evident that grouping these cancers as one type would not allow clinically useful biomarkers to be identified. Most of the patients used in this study are oral cancer patients. If so, recurrence is lower as opposed to HPV-positive. Please justify the selection of the sample cohort.

   We agree that HNSCC is a heterogeneous disease and differs, biologically and clinically, based on the anatomical site of origin. But it is important to note that all HNSCC are epithelial tumors. In this context, one of the most relevant findings of our study relies in the fact that the microRNAs selected here are epithelial markers and they seem to be able to detect the presence of metastatic cells in the lymph nodes from HNSCC patients in general, regardless of tumor origin.

   Actually, the FFPE tissue samples analyzed in the study were all from the oral cavity, more specifically, from the lower part of the oral cavity (Supplementary Table 1). We chose these tumor sites because previous studies had pointed out an increased risk of subclinical disease in lymph nodes from patients with oral tongue and floor of the mouth tumors.
(Elsheikh et al. 2006). Of note, the best clinical management of the neck of patients with initial tumors at these locations is controversial and, therefore, a molecular markers that increases the accuracy of the detection methods for subclinical disease in the lymph nodes would help in the choice of best approach to treat the neck of the HNSCC patients.

Noteworthy, the two microRNAs validated as markers for the presence of metastatic cell in the lymph nodes from patients with tumors the oral cavity were also evaluated in a group of FNA samples collected from patients harboring tumors in different stages occurring in different topologies (oral cavity, pharynx and larynx (Supplementary Table 1). These results showed that these markers can successfully detect metastatic lymph nodes, regardless of the tumor site.

These markers showed to be able to detect, with high accuracy, the presence of metastatic lymph nodes in patients with tumors in oral tongue, a tumor site epidemiologically recognized as rarely associated with HPV infections, as well as, in tumor of oropharynx, tumors frequently correlated to HPV. Based on this, we strongly believe that HPV status will not influence the detection of the presence of epithelial cells in the lymph nodes by the markers selected here.


3. Studies show that 10-40% of HNSCC patients with histopathologically negative neck lymph nodes eventually develop regional metastases, suggesting that metastatic cells present in the lymph nodes could not be detected at diagnosis Please specify which type of HNSCC? HPV +ve or negative. Recurrence both local and distant metastasis is higher in HPV+ve cancers.
This affirmation is based on the reports of 3 independent studies. Two of them (Rhee et al., 2002; Becker et al., 2004) revised clinical data of HNSCC (multiple sites were analyzed) and the third one revised data from supraglottic carcinomas (Tu, 1999). Unfortunately, none of them reported the HPV status of the cases analysed.


4. Under statistical analysis please explain whether the data set is normally distributed. If that is the case you can use T test. If not you need to use non-parametric analysis. This section needs to be justified

DataAssist software, recommended by the PCR array supplier to be used in the analysis of PCR-array data, does not give the option of performing a non-parametric test. This software only allows conducting a two-sample two-tailed Student’s t-test to compare ΔCT values of the two groups. So, influenced by the supplier recommendation, the statistic test available in the software we initially applied t-test in the screening of our discovery set. We perfectly agree that given the small number of samples evaluated in the discovery set, a normal distribution could not be reached and the use of a parametric test may be too optimistic (too low p-value). The analysis using the Shapiro-Wilk test confirmed the data set is not normally distributed. Therefore, the entire PCR-array data was re-analyzed using a non-parametric test (Rank Products). According to this new analysis, 61 microRNAs were selected because they presented at least 2-fold upregulation in all four metastatic samples.
with a $p<0.05$ (Rank Products). Noteworthy, even with the new statistical analysis, the main findings of the manuscript were confirmed.

5. Mean Ct values of U6 and U47 small nuclear RNAs were used for normalization, and the mean Ct of 10 non-metastatic lymph nodes was used as reference. Please clarify ct of 10? is this the threshold? What was the average Ct for the house keeping genes between FNA and FFPE samples.

As mentioned above, the $2^{-\Delta\Delta Ct}$ method used to calculate fold-change values requires a control group to be used as reference. Thus the fold-change expression of a given “marker” in cases samples will always be compared to the expression of the same “marker” in the reference group (relative expression). Therefore, we randomly selected 10 samples from patients whose all lymph nodes resected were negative for metastases (according to the step-sections and IHC for cytokeratins) and use them as reference group in the $2^{-\Delta\Delta Ct}$ analyses.

This statement was better clarified in the revised version of the manuscript.

6. Reviewer found it very difficult to follow the study design. Initially the authors undertook a discovery study where by FFPE samples from LN positive and negative patients were compared. Followed by a retrospective study whereby the authors collected FNA and Ln from both ln+ve and Ln-ve patients. Is this the case?

First of all, it is important to mention that this study only evaluated metastatic and non-metastatic lymph node samples, no primary tumor tissues were assessed. The discovery set comprised metastatic and non-metastatic FFPE lymph nodes from patients with tumor in the oral cavity. This analysis allowed us to select 8 microRNAs to be evaluated in the validation set. The samples evaluated in the validation set included an additional group of metastatic and non-metastatic FFPE lymph nodes from patients harboring oral cavity
tumors. This analysis identified 2 microRNAs with high accuracy, sensitivity and specificity. Later, these two markers were evaluated in a prospective cohort comprised by FNA samples collected from negative and positive lymph nodes from HNSCC samples (oral cavity, pharynx and larynx).

7. Supplementary table 2 indicates that: Clinical features of the 79 HNSCC patients enrolled in the FNA validation data set – 113 FNA biopsies were collected from patients? How come initially 79 and then 113?

As indicated in the results section “Identification of metastatic cell deposits in FFPE and FNA lymph node samples” of the manuscript, “whenever possible, FNA biopsies were conducted in macroscopically positive and negative lymph nodes from the same patient”. So, although only 79 patients were enrolled, sometimes more than one FNA biopsies were collected from the same patient (whenever possible), accounting for 113 FNA biopsies collected. This is why there is a difference between the number of patients enrolled (79) and the biopsies collected (113).

This information was added to the legend of the Supplementary Table 2.

8. Explain the youden index and how this was calculated?

The receiver operating characteristic (ROC) curve is used to evaluate the effectiveness of a certain biomarker in the determination of a diseased and non-diseased population. The ROC curve is a plot of sensitivity (Se) versus 1-specificity (1-Sp) at all possible cutoff (c) levels. The Youden Index (J) is a way to summarize ROC curve statistics in the interpretation and evaluation of a biomarker. It defines the maximum potential effectiveness of a biomarker and can be formally defined as $J = \max_c \{Se(c) + Sp(c) − 1\}$ (Ruopp et al., 2008). The index is represented graphically as the height above the chance line (diagonal line in the ROC curve). The maximum value of the Youden's index is represented by the cutoff.
value with the highest height above the chance line (Powers, 2011). In the present study, the Youden index was obtained from the coordinate table of the ROC curves as the value in which the difference between sensitivity and 1-specificity is maximum.

This point was clarified in the revised version of the manuscript.


9. Under method section also includes the RNA purity and integrity. It is difficult to isolate intact good RNA from FFPE slides. It is worth while letting the readers know what strategies were employed to yield high RNA. Also, there is no mention of DNase treatment? Why> Is this because the procedure that you follow selectively remove large oligonucleotides?

For the extraction of RNA from the FFPE samples, the Recoverall Total Nucleic Acid Isolation kit (Ambion) was used according to the manufacturer’s recommendations. One of the last steps in the Kit protocol is a treatment with DNAse.

Due to the scarcity of RNA quantity recovery from the FFPE samples, no step for evaluation of the RNA purity or integrity was done. We just went through the PCR analysis and used the result of the qPCR reaction with the internal controls (U6 and U47) to infer the quality of the RNA sample. In other words, if we obtained a positive amplification with the internal controls, the sample was considered “good” and kept in the study. Conversely, if no amplification was observed, the sample was discarded. It is important to mention that all the samples included in the study showed amplification of both internal controls.
10. **Minor comments: Discussion is too long. Needs to be focused.**

*This was addressed in the revised version of the manuscript.*