

# SCIENTIFIC REPORTS



OPEN

## Insights into immune responses in oral cancer through proteomic analysis of saliva and salivary extracellular vesicles

Received: 04 March 2015  
Accepted: 09 October 2015  
Published: 05 November 2015

Flavia V. Winck<sup>1</sup>, Ana Carolina Prado Ribeiro<sup>2</sup>, Romênia Ramos Domingues<sup>1</sup>, Liu Yi Ling<sup>3</sup>, Diego Mauricio Riaño-Pachón<sup>3</sup>, César Rivera<sup>1,5</sup>, Thaís Bianca Brandão<sup>2</sup>, Adriele Ferreira Gouvea<sup>2</sup>, Alan Roger Santos-Silva<sup>4</sup>, Ricardo D. Coletta<sup>4</sup> & Adriana F. Paes Leme<sup>1</sup>

The development and progression of oral cavity squamous cell carcinoma (OSCC) involves complex cellular mechanisms that contribute to the low five-year survival rate of approximately 20% among diagnosed patients. However, the biological processes essential to tumor progression are not completely understood. Therefore, detecting alterations in the salivary proteome may assist in elucidating the cellular mechanisms modulated in OSCC and improve the clinical prognosis of the disease. The proteome of whole saliva and salivary extracellular vesicles (EVs) from patients with OSCC and healthy individuals were analyzed by LC-MS/MS and label-free protein quantification. Proteome data analysis was performed using statistical, machine learning and feature selection methods with additional functional annotation. Biological processes related to immune responses, peptidase inhibitor activity, iron coordination and protease binding were overrepresented in the group of differentially expressed proteins. Proteins related to the inflammatory system, transport of metals and cellular growth and proliferation were identified in the proteome of salivary EVs. The proteomics data were robust and could classify OSCC with 90% accuracy. The saliva proteome analysis revealed that immune processes are related to the presence of OSCC and indicate that proteomics data can contribute to determining OSCC prognosis.

Oral cancer represents 1 to 2% of all types of cancer worldwide, and oral squamous cell carcinoma (OSCC) is the most frequent histopathological type reported among patients<sup>1</sup>. This disease can affect different sites in the oral cavity, such as the tongue, floor of the mouth and cheeks, and smoking and alcohol ingestion is responsible for approximately 90% of OSCC cases<sup>2–4</sup>. Half of the patients with oral cancer are diagnosed only when the disease has reached an advanced clinical stage (III/IV), leading to a five-year survival rate of only 20% in these cases<sup>1,5,6</sup>. OSCC prognosis is currently based on the clinical staging system of tumor-lymph node-metastasis of the disease (TNM system). However, this system is not optimal because tumors may present distinct biological characteristics even at similar developmental stages<sup>7</sup>. Thus, the identification of additional parameters or biological markers that assist with determining the prognosis of patients with OSCC is essential.

<sup>1</sup>Laboratório de Espectrometria de Massas, Laboratório Nacional de Biociências, LNBio, CNPEM, Campinas, SP, Brazil. <sup>2</sup>Instituto do Câncer do Estado de São Paulo, Octavio Frias de Oliveira, ICESP, São Paulo, SP, Brazil. <sup>3</sup>Laboratório Nacional de Ciência e Tecnologia do Bioetanol, CTBE, CNPEM, Campinas, SP, Brazil. <sup>4</sup>Faculdade de Odontologia de Piracicaba, Universidade Estadual de Campinas, UNICAMP, Piracicaba, SP, Brazil. <sup>5</sup>Departamento de Ciencias Básicas Biomédicas, Universidad de Talca (UTALCA), Talca, Chile. Correspondence and requests for materials should be addressed to A.F.P.L. (email: adriana.paesleme@lnbio.cnpem.br)

One promising strategy for the discovery of new biomarkers consists of the identification of the protein profile of body fluids, such as saliva, which may be used to characterize a specific disease. The repertoire of disease-related proteins and other molecules can be identified by mass spectrometry. Although the use of saliva is not novel, this approach is receiving great interest as a diagnostic fluid because harvesting saliva is performed using non-invasive methods. Previous studies reporting the proteomic analysis of saliva from healthy individuals have already indicated the potential of this approach for monitoring health<sup>8</sup>. Furthermore, recent studies have shown that saliva contains many signaling molecules that may be indicative of cancer<sup>9,10</sup>.

Moreover, cancer cells can generate several types of extracellular vesicles, ranging from 30 nm to a few micrometers in diameter, which may be shed from the cytoplasmic membrane<sup>11</sup>. These extracellular vesicles (EVs) can deliver molecules (such as proteins, mRNA, microRNA, rRNA, tRNA, DNA and lipids) that have been suggested to participate in important intracellular signaling mechanisms even in distant target cells<sup>11–13</sup>. Although many potential biomarkers for oral cancer have been identified in human saliva, the role of such molecules in oral cancer is not completely understood<sup>10,14–16</sup>. Several proteins, including CD44, IL-6, IL-8 and defensin-1, have been suggested to be OSCC biomarkers. The expression of these molecules has also been reported in other types of cancer, indicating that common cancer cell responses may underlie tumor progression<sup>17–20</sup>. Furthermore, global similarities between the differing types of cancer indicate that cancer complexity may reside in the cellular host responses, which could indeed influence the mechanisms by which cancer progresses in different individuals.

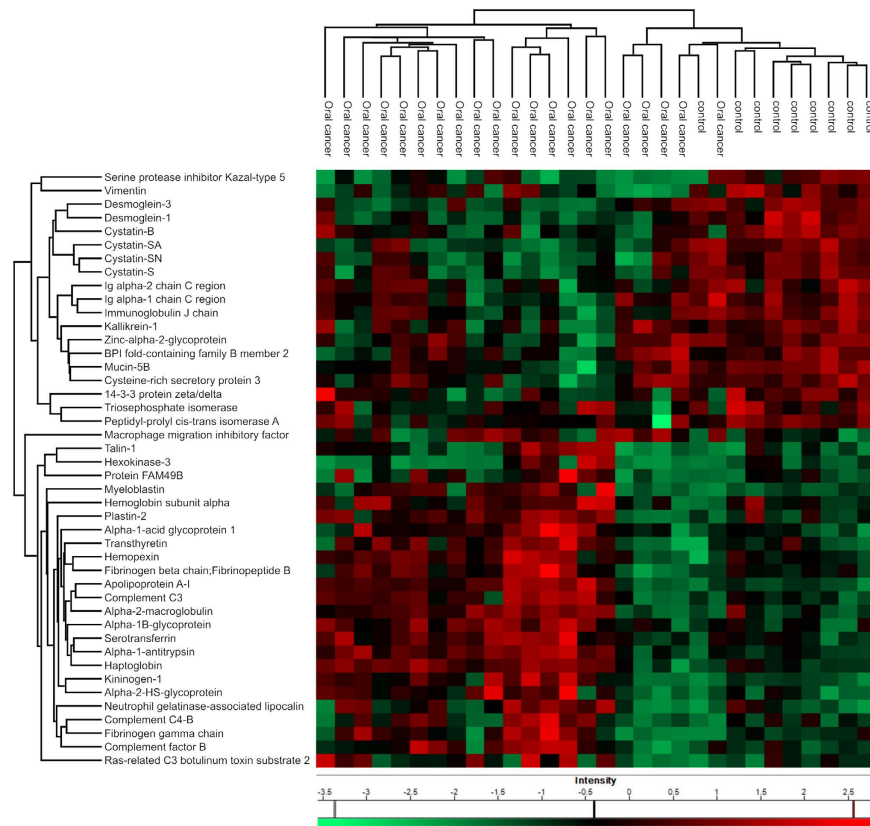
In the present study, we examined the proteome profile of whole saliva and salivary EVs from patients with oral cancer (patients with and without tumor lesions) and healthy individuals by mass spectrometry analysis. Due to significant differences in processes related to inflammatory and humoral immune responses, peptidase inhibitor activity, iron coordination and protease binding, the two classes of individuals (healthy vs. OSCC) were distinguished with 90% accuracy based on the proteomics data. Moreover, proteome functional annotation revealed that differentially expressed salivary proteins may indeed be related to cell-to-cell signaling and cellular interaction mechanisms.

## Results

**Differentially expressed proteins in whole saliva from oral cancer patients showed potential roles in peptidase regulation and immune responses.** We analyzed the whole salivary proteome and the proteome from salivary EVs isolated from healthy individuals and oral cancer patients (“lesion” plus “no lesion” individuals). The clinical classifications of the patients are listed in Supplemental Table 1. Relative protein quantification was performed using the label-free quantification (LFQ) method, and our data showed high reproducibility (Supplemental Figure 1 and Supplemental Figure 2). This analysis resulted in the identification of 507 proteins after data pre-processing (excluding contaminants, reverse sequences and only identified by site) (Supplemental Table 2) from which 147 proteins had at least five valid LFQ intensity values in at least one group (healthy,  $n = 9$ , or oral cancer,  $n = 21$ ). The filtered dataset containing 147 proteins was subjected to further data analysis using a one-way ANOVA test ( $p < 0.05$ ) resulting in 44 proteins showing differential expression. In addition, ANOVA  $p$ -values corrected by Storey’s method are shown (Fig. 1, Supplemental Table 3). Most of these differentially expressed proteins showed reduced expression in oral cancer patients. Additionally, nineteen proteins were exclusively identified in the healthy group, and 247 proteins were exclusively identified in the oral cancer group.

The results of GO enrichment analysis of the differentially expressed proteins (44 proteins, ANOVA  $p < 0.05$ ) showed a significant overrepresentation of biological processes related to acute inflammatory response ( $p$ -value =  $3.3E-07$ ), regulation of humoral response ( $p$ -value =  $3.1E-06$ ) and regulation of hydrogen peroxide metabolic processes ( $p$ -value =  $1.67E-04$ ). Cellular components related to cytoplasmic membrane-bounded vesicle lumen ( $p$ -value =  $5.9E-10$ ) and endocytic vesicle lumen ( $p$ -value =  $6.5E-06$ ) components were overrepresented in the dataset. According to the IPA (Ingenuity® Systems, <http://www.ingenuity.com>) analysis, the molecular function cell-to-cell signaling and interaction was overrepresented in the dataset of differentially expressed proteins (Supplemental Table 4). Furthermore, overrepresented molecular function GO terms included humoral immune responses, peptidase inhibitor activity and protease binding ( $p$ -value  $< 0.05$ ). A network of the overrepresented GO terms is shown in Fig. 2.

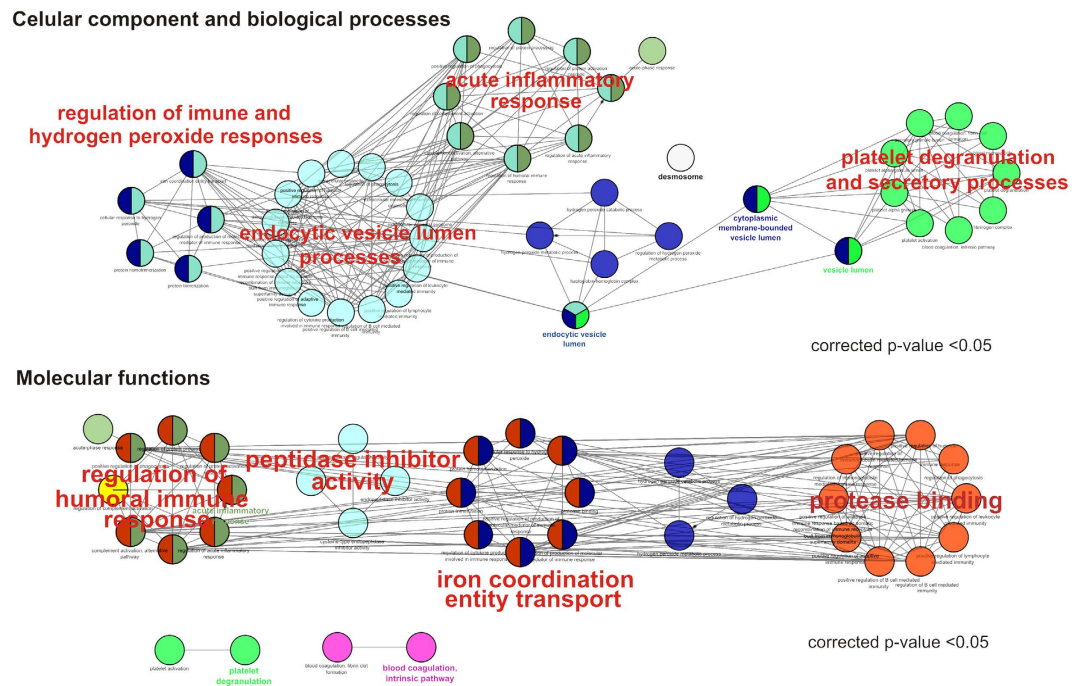
**Proteomic data analysis of whole saliva from healthy and oral cancer patients permitted classification of patients.** Principal component analysis (PCA) was performed using the proteome LFQ data (no missing values) from whole saliva and indicated that groups of healthy individuals can be distinguished from the oral cancer individuals (Supplemental Figure 3). Therefore, we applied the support vector machines (SVM) method to the proteome dataset to evaluate whether the dataset could classify the two groups of individuals. Furthermore, proteins that could classify the two groups of individuals (healthy and OSCC) were identified through the application of the feature selection method. For the generation of the classification model using SVM, 571 attributes (protein intensity normalized by  $z$ -score) and 30 instances (individuals) were used, and a 3-fold cross-validation method was performed. Our classification model resulted in high precision (0.925), recall (0.9), and ROC (0.929) and a Kappa value of 0.78. The proteins identified as the best features for data classification are listed in Table 1 with their functional annotations.



**Figure 1. Differentially expressed proteins in whole saliva from human oral cancer.** LFQ intensity values of the salivary proteins identified from oral cancer patients were compared to the LFQ values of salivary proteins identified from healthy individuals, and proteins with a significant differential abundance were detected using the ANOVA ( $p < 0.05$ ) method. Clustering analysis of the data was performed using a Euclidian distance method for the z-score normalized LFQ values. The z-score normalized values are shown in a heat map that represents the variation in protein abundance between the analyzed samples. Groups of oral cancer and healthy individuals and the corresponding protein names are indicated.

**Oral cancer patients with lesions and without tumor lesions have differing proteome repertoires.** The influence of the presence or absence of tumor lesions in patients diagnosed with oral cancer was investigated. Among the total 507 proteins identified, 13 proteins were identified as differentially expressed between the “lesion” ( $n = 12$ ) and “no lesion” ( $n = 9$ ) groups (Fig. 3). A group of 38 proteins were identified exclusively in the saliva of patients with lesions, and 5 proteins were exclusively identified in the group of patients without lesions, considering proteins identified in at least two individuals. According to IPA knowledge base information, some of the differentially expressed proteins may play roles in processes related to cell movement, morphology and cell death (Supplemental Table 5).

**Proteomic profiling of salivary extracellular vesicles indicated overrepresentation of antigen binding and enzyme inhibitory functions.** The size distribution of salivary EV samples was measured using the NTA, revealing the presence of EVs with an average size of 177.19 nm ( $SD = 17.22$  nm) (Fig. 4A). Western blotting analysis for flotilin-1 indicated a greater abundance of this marker protein in the isolates of EVs (Fig. 4B). Furthermore, the structure of isolated EVs was observed by transmission electron microscopy (TEM-HR), which revealed that the EVs were intact after isolation (Fig. 4C). A total of 381 proteins were identified by mass spectrometry in the EV proteome from the two groups of individuals (healthy ( $n = 10$ ) and oral cancer ( $n = 7$ )) (Supplemental Table 6). Considering proteins that showed a minimum of three valid expression values in at least one group, 139 proteins were subjected to statistical analysis. Eight proteins from the EVs were differentially expressed (ANOVA,  $p < 0.05$ ) between the two groups (Fig. 5). Among the total 381 proteins identified, 18 proteins were exclusively detected in the healthy group and 4 proteins were exclusively identified in the oral cancer group, considering proteins identified in at least two individuals. GO enrichment analysis revealed that the biological processes “antigen binding,” “enzyme inhibitor” and “endopeptidase regulatory functions” were overrepresented in the EVs proteome (Supplemental Figure 4, Supplemental Table 7).



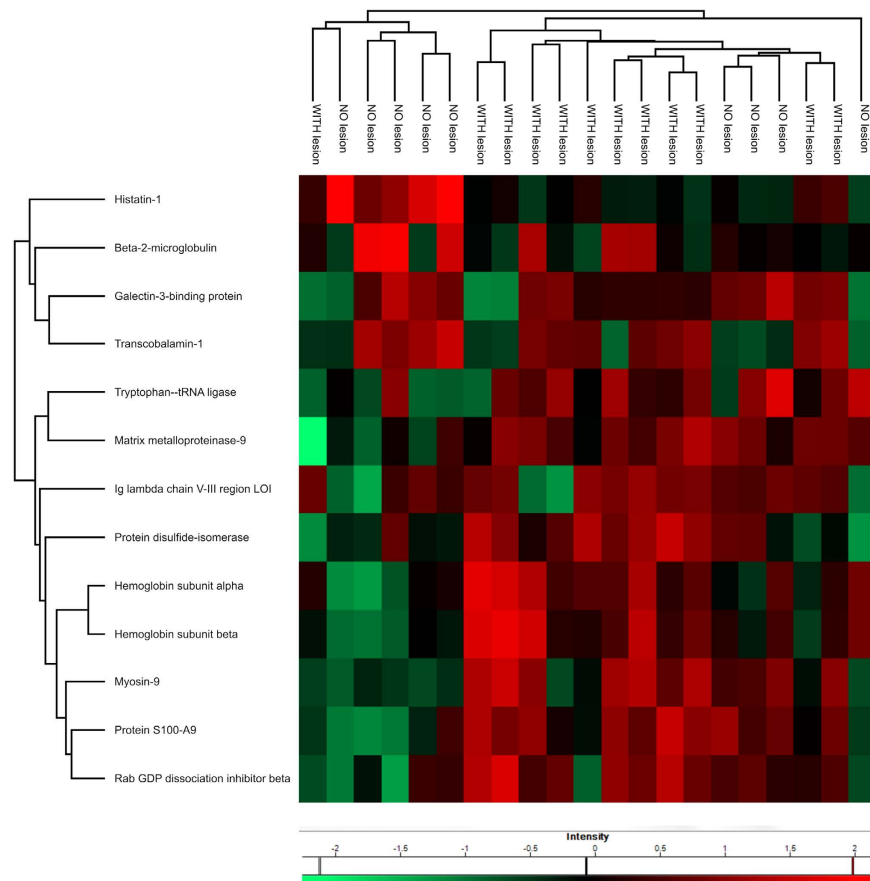
**Figure 2. Network of overrepresented GO terms in the dataset of differentially expressed proteins in the whole saliva proteome.** Proteins identified as differentially expressed in the whole saliva of patients with oral cancer and healthy individuals were subjected to a GO term enrichment analysis using the ClueGO plugin. Overrepresented GO terms ( $p$ -value  $< 0.05$ ) for Biological Processes, Cellular Components and Molecular Function categories were visualized using the Cytoscape suite. The most significant GO terms within the dataset of significantly enriched terms are indicated in red font.

UniProt_AC	Gene names	Protein names	Differentially expressed* (Yes/No)
P01034	CST3	Cystatin-C	No
P23280	CA6	Carbonic anhydrase 6	No
P32926	DSG3	Desmoglein-3	Yes
Q02818	NUCB1	Nucleobindin-1	No
Q8TDL5	BPIFB1	BPI fold-containing family B member 1	Yes
Q9HC84	MUC5B	Mucin-5B	Yes
Q9NQ38-3	SPINK5	Serine protease inhibitor Kazal-type 5	Yes
P20591	MX1	Interferon-induced GTP-binding protein Mx1	No

**Table 1. Proteins of whole saliva with potential for classification of oral cancer individuals were selected based on feature and attribute selection methods.** \*Proteins significantly differentially expressed in the whole salivary proteome or in EVs salivary proteome analysis.

**Proteomics data revealed prognosis-associated proteins and functional connections among protein-coding genes and drugs.** The intensity values of the differentially expressed proteins identified in our study were used to analyze survival time. We observed that among the differentially expressed proteins, only the protein peptidyl-prolyl cis-trans isomerase A (PPIA, also known as cyclophilin-A) was statistically significant in the analysis of the mean survival time (Fig. 6). Notably, reduced abundance of PPIA appears to be a factor that may predict poor prognosis of OSCC patients. Upon observing that differentially expressed proteins may indeed possess prognostic value for patients with OSCC, we further investigated functional connections associated with these proteins.

Using the cMAP database<sup>21</sup>, eight compounds were identified as significantly related to alterations in the expression of the genes that code for the differentially expressed proteins (Supplemental Table 8).



**Figure 3. Differentially expressed proteins in response to the presence or absence of tumors in patients diagnosed with oral cancer.** LFQ intensity values ( $\log_2$ ) of the salivary proteins identified from diagnosed oral cancer patients with tumors (with lesion) were compared to patients without tumors (no lesion). Differentially expressed proteins were detected using the ANOVA method ( $p < 0.05$ ). A Euclidean distance clustering analysis of these data was performed, and the differential protein abundance is shown in the heat map of normalized z-score values.

The small compounds terazosin, 15-delta prostaglandin J2, PNU-020031, acyclovir, dextromethorphan, yohimbic acid, benzthiazide and pseudopelletierine were ranked as the most significant compounds, with  $p$ -values  $< 0.01$ .

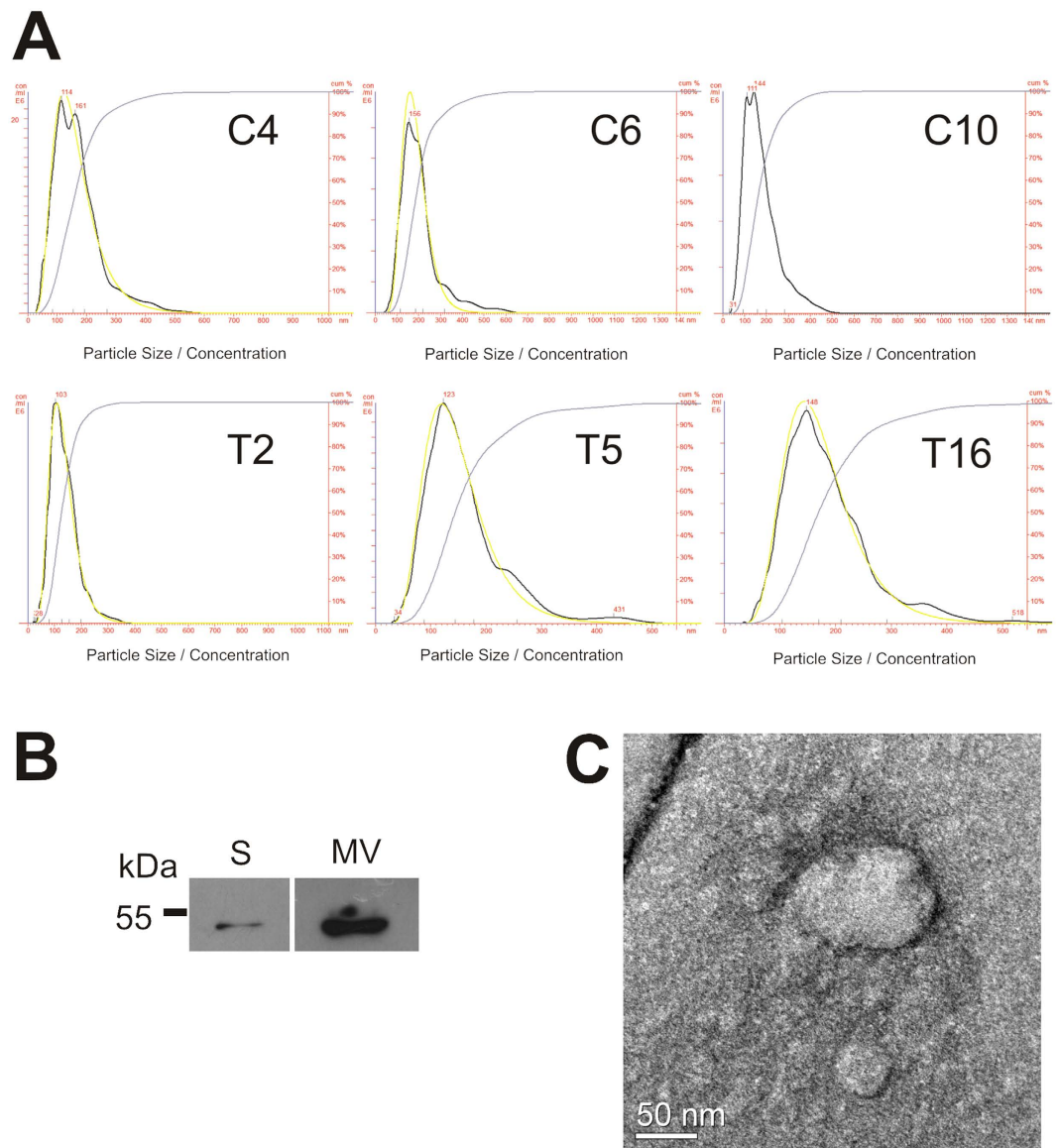
Furthermore, a cross-cancer analysis of the frequency of genetic alterations (mutation, deletion, amplification and multiple alterations) in the genes that code for the differentially expressed proteins in the saliva of OSCC patients indicated that in general, mutations in the genes coding for the differentially expressed proteins occur in less than 30% of the cases of head and neck cancer in datasets reported through the cBioPortal for cancer genomics (Supplemental Figure 5a). The frequency of alterations in the genes coding for proteins reported as classificatory for OSCC through the feature selection method is less than 15% (Supplemental Figure 5b) according to the information available from cBioPortal.

Furthermore, S-score ([www.bioinformatics-brazil.org/S-score/](http://www.bioinformatics-brazil.org/S-score/)) classification analysis of the list of differentially expressed proteins and proteins that were robust for classification of OSCC indicated that depending on the types of cancer compared, the same proteins may play roles as tumor oncogenes or tumor suppressors. The S-score calculation revealed that the genes coding for the differentially expressed proteins in the salivary proteome of OSCC patients display expression patterns that are most similar to the gene expression profile of the ovarian cancer dataset (Supplemental Figure 6).

## Discussion

The shotgun proteome analysis of saliva from healthy and oral cancer individuals revealed the identity of hundreds of proteins that support previous studies of the intricate effects of immune-response mechanisms in tumor development. Statistical analysis of the proteomics data revealed the modulation of expression levels of proteins related to networks and pathways that control immune responses in patients with oral cancer (Fig. 2).

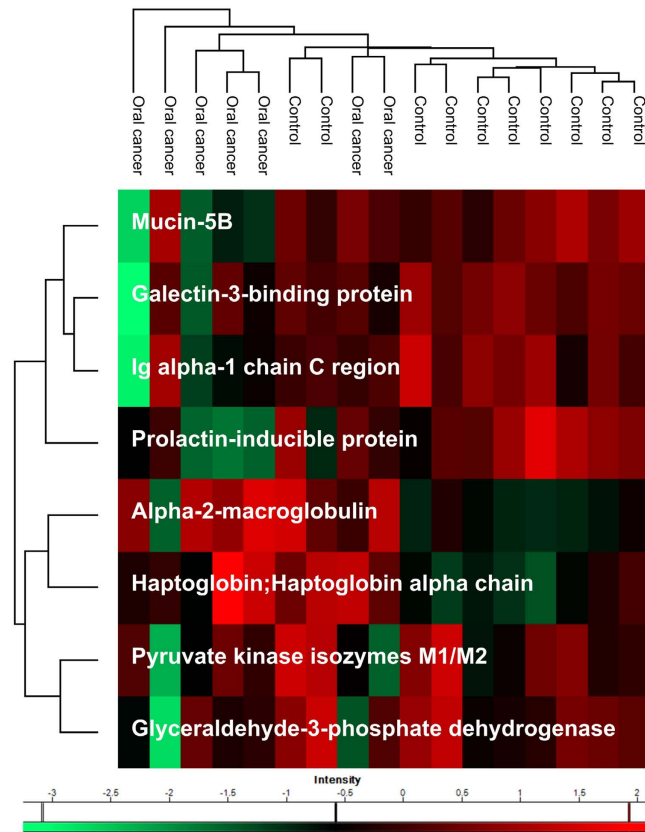
Jessie and colleagues have explored the aberrant expression of proteins in the saliva of OSCC patients<sup>15</sup>, and the proteins serotransferrin, hemopexin, alpha-1beta-glycoprotein, alpha-1-antitrypsin, haptoglobin,



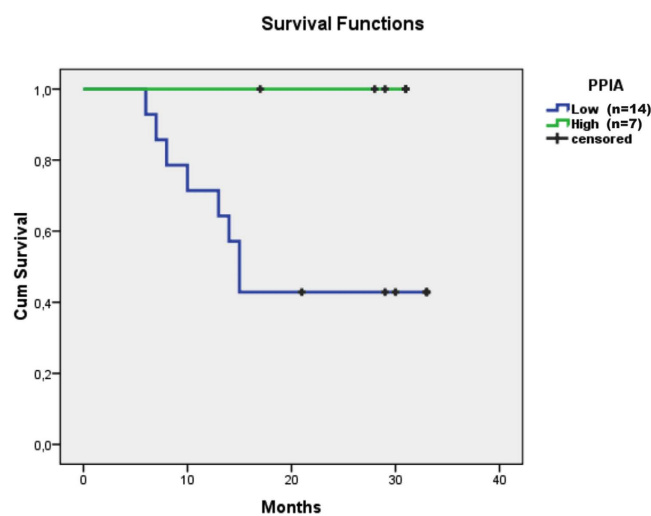
**Figure 4. Extracellular vesicles isolated from saliva.** Extracellular vesicles were isolated from the saliva of healthy and oral cancer patients using ultracentrifugation. (A) Nanoparticle tracking analysis of the extracellular vesicles isolated from healthy individuals (C4, C6, C10) and oral cancer patients (T2, T5, T16); (B) Immunochemical detection of the protein flotilin-1, a protein marker of extracellular vesicles, in the whole saliva proteome (S) and in the proteome of extracellular vesicles (EV); (C) Transmission electron microscopy of isolated extracellular vesicles using a negative staining method.

fibrinogen  $\beta$ , complement C3 and transthyretin were detected in the saliva of OSCC patients<sup>15</sup>, in agreement with our proteomics results presented here. Of note, Jessie performed a very similar protein extraction procedure to that used in our present study, which is important for the establishment of a panel of biomarkers in biofluids. Haptoglobin and alpha-1-antitrypsin have previously been associated with OSCC in tissue samples and are known components of the acute-phase response, generally produced in the liver in response to the effects of cytokines<sup>15,22</sup>. However, studies of blood serum samples have indicated that acute-phase proteins, including complement C3 and haptoglobin, correlate with the progression of breast cancer, small cell lung carcinoma and lung adenocarcinoma<sup>23</sup>.

Previous proteomic analysis of human saliva from healthy individuals has revealed that biological processes associated with metabolic and regulatory pathways are overrepresented<sup>24</sup>, with nearly 50% of the proteins functionally associated with metabolism and energy pathways, protein metabolism and cell communication<sup>16</sup>. Therefore, the biological processes overrepresented in our dataset of salivary proteins from OSCC patients may indeed represent specific cellular responses and do not represent common features of the healthy saliva proteome.



**Figure 5. Proteins from extracellular vesicles with a significant differential abundance between healthy and oral cancer groups.** Extracellular vesicles were isolated from the saliva of healthy and oral cancer individuals, and the EV protein content was analyzed by shotgun proteomics. Proteins with a significant differential abundance between healthy and oral cancer groups were identified using the label-free quantification (LFQ) method (ANOVA,  $p < 0.05$ ). The Euclidean distance clustering method was used to visualize the variations in the protein abundance between the samples and is shown as normalized z-score log<sub>2</sub> LFQ values in a heat map.



**Figure 6. PPIA survival function from lifespan data.** Kaplan-Meier curves demonstrate that patients with a PPIA low profile succumbed to the disease much earlier than those with a high profile. After 1.5 years, approximately 40% of the low profile patients survived, but all subjects in the high intensity group were alive at follow-up. A log-rank test shows significant differences between pairs, with  $p$ -value  $< 0.05$ . The number of individuals included in the analysis is shown in the figure legend.

Moreover, knowledge available on the cBIOPortal indicated that most protein coding genes for the differentially expressed proteins in the OSCC phenotype have a genetic mutation frequency of less than 30% in head and neck tumors. Furthermore, the results of the S-score analysis performed for the same dataset of protein coding genes (Supplemental Figure 6a) and proteins that may classify OSCC (Supplemental Figure 6b) indicated that complete agreement does not exist between the gene expression patterns of these genes in different types of tumors. No clear indication exists regarding the role of these genes in OSCC as putative oncogenes or tumor suppressors. However, the expression pattern (up- or down-regulation) of the differentially expressed protein coding genes in OSCC matched the gene expression pattern from six putative oncogenes and suppressor genes in an ovarian cancer dataset. Notably, tumorigenesis in ovarian carcinoma occurs through an epithelial-to-mesenchymal transition, in a manner similar to the tumorigenesis observed in OSCC<sup>25</sup>.

The high abundance of the proteins complement C3, complement factor B (CFB) and complement C4-B in the saliva of OSCC patients (Supplemental Table 2) has also previously been observed in tissue samples<sup>16</sup>. Complement activation has recently been shown to occur in ovarian cancer cells due to an autocrine effect that produces complement C3 protein, which leads to alterations in the tumor microenvironment by increasing the number of myeloid-derived suppressor cells and reducing the number of cytotoxic T cells that infiltrate into the tumor, which promotes tumor growth<sup>26</sup>. According to our signaling network analysis, complement C3 may be autoactivated and bind to the CFB protein, inducing the activation of the complement system (Supplemental Figure 7). Complement C3 has also been suggested to contribute to the secretion of prostaglandin E2 (PGE2) in aberrant immune responses during induced inflammation in mammals<sup>27</sup>. Furthermore, the overexpression of macrophage migration inhibitory factor (MIF) in the saliva from OSCC patients may also correlate with inflammatory responses through alterations in cell movement, adhesion, proliferation and invasion in head and neck squamous cell carcinoma<sup>28</sup> via its functional role activating CD74 protein<sup>29</sup>. As a general activator of the immune response, MIF can overstimulate the secretion of prostaglandin E2 (PGE2) in endometriotic stromal cells<sup>30</sup> and has direct and indirect effects on tissue remodeling and angiogenesis<sup>31</sup>.

Furthermore, the proteome of salivary EVs displays overrepresentation of proteins related to specialized vesicular structures secreted by the cells, including “membrane-bound vesicle” (corrected p-value = 5.44E-11), “extracellular organelle” (corrected p-value = 4.23E-10), “extracellular membrane-bound organelle” (corrected p-value = 4.23E-10) and “extracellular vesicular exosome” (corrected p-value = 5.43E-10). GO biological processes overrepresented in the proteome of EVs included “leukocyte migration involved in inflammatory response” (corrected p-value = 0.0011) and “chemokine biosynthetic process” (corrected p-value = 0.0031). Molecular functions related to “lipoprotein particle receptor binding” (corrected p-value = 0.0030), “peroxidase activity” (corrected p-value = 0.0032) and “oxidoreductase activity” (corrected p-value = 0.0032) are overrepresented in the proteome of EVs (Supplemental Figure 4, Supplemental Table 7). Interestingly, the proteome of EVs isolated from the OSCC patients was enriched in proteins with molecular and cellular functions related to “molecular transport,” mainly comprising transport of metals, and “cellular growth and proliferation” (Supplemental Table 9). These results indicate that salivary EVs may transport signaling molecules in the tumor microenvironment and may attract inflammatory cells to the tumor site by chemotactic mechanisms, which may be important in the subsequent immunoeediting of inflammatory cells<sup>32</sup>. The study of the function of these molecules secreted by inflammatory cells may better explain how cancer cells can benefit from the function of the inflammatory system. Furthermore, tumor-derived EVs may be related to the formation of a pre-metastatic niche and signaling events in cell-to-cell communication<sup>33,34</sup> through feedback loops between inflammation and tumor growth, possibly involving proteins such as STAT3 and IL-6R in the processes of cancer invasion and metastasis<sup>35</sup>.

The salivary proteome analysis from OSCC patients with and without lesions indicated that patients with active oral malignant lesions have a higher abundance of proteins related to a highly invasive phenotype and an immune response (Fig. 3, Supplemental Table 10). Among these proteins, we identified matrix metalloproteinase-9 (MMP-9), protein S100A9, myosin-9 (NMMHC II-a) and Rab GDP dissociation inhibitor beta (Supplemental Figure 8). Overexpression of MMP-9 has been detected in saliva of OSCC patients and has been implicated in promoting the invasive behavior of colon cancer cells<sup>14,36</sup>, pathological neovascularization and increased cell proliferation in other types of cancer<sup>37</sup>.

Furthermore, the protein S100A9, which functions as a heterodimer with S100A8, is related to the regulation of immune responses and increases in cell proliferation and motility in multiple cancers<sup>38</sup>. Overexpression of S100A9 has been shown to increase the size and volume of skin tumors<sup>39</sup> and has also been related to chemo-attraction of leukocytes and macrophages, which may trigger pro-inflammatory responses in the tumor microenvironment<sup>38</sup>. However, more recent evidence has demonstrated that the S100A9/S100A8 heterodimer may actually function as a suppressor of acute inflammatory responses and contribute to reducing the gene expression of IL6, IL1-beta and TNF alpha<sup>40</sup>.

Our results indicate that PPIA protein is less abundant in the whole saliva of OSCC patients and is the top performing prognostic candidate protein; however, further validation in a larger population is ongoing. Previous studies have suggested that PPIA plays important roles in several mechanisms, including protein folding, trafficking and T-cell activation<sup>41</sup>.

We also observed modulation of proteins related to the production of vesicles and cytoskeleton dynamics. The high abundance of the Rab GDP dissociation inhibitor beta (Rab GDI) protein in the



OSCC salivary proteome may indicate an increased need by cancer cells for the generation and secretion of EVs; Rab GDI has been suggested to be an important element in the control of vesicle trafficking<sup>42</sup>.

Furthermore, overexpression of the myosin-9 protein in patients with tumor lesions reinforces the suggested role of myosin-9 in cancer progression through affecting cell migration and actin cytoskeleton remodeling by interacting with the protein Arl13b<sup>43</sup>. Additional evidence has indicated that myosin-9 is involved in the directional migration of immune cells, including neutrophils<sup>44</sup>, and movement of T lymphocytes through the endothelial junctions<sup>45</sup>. In breast cancer cells, functional inhibition of myosin-9 results in impaired cell invasion<sup>46</sup>. In our study, myosin-9 was identified in 14 of the 17 EV samples, indicating that myosin-9 may be important in the regulation of the dynamics and organization of the cytoskeleton structures of the EVs secreted by normal and cancer cells.

The fact that OSCC individuals without tumor lesions still exhibit a salivary proteomic repertoire that differs from the healthy group is an indicator that the host response continues to be at least partially modulated once the individuals develop cancer or in response to the medical procedures performed in non-lesion patients.

Our results indicate that differentially expressed salivary proteins might be related to humoral immune responses, which in turn may be important for oral tumor development and indicate tumor progression. The proteome of salivary EVs also displayed the presence of proteins with functions related to an inflammatory response in addition to transport of metals and cellular growth and proliferation.

Therefore, the proteomic analysis of whole saliva from healthy individuals and OSCC patients indicated that the composition of saliva may indeed reflect the response of individuals diagnosed with oral cancer. The salivary composition of the patients diagnosed with OSCC is different from the saliva of healthy individuals, independently of the presence of the tumor itself, given that the saliva of patients without lesions also exhibited overexpression of known cancer-related proteins.

However, the mechanisms by which the saliva components and the immune cells interact to modulate the cellular responses in the cancer cells is not completely understood. An understanding of the role of extracellular vesicles in this cell-to-cell communication may reveal the mechanisms by which immune responses and inflammatory cells contribute to positively affect tumor growth.

## Methods

**Saliva collection and preparation.** The study was approved by the ethics review board of the Instituto do Câncer do Estado de São Paulo (ICESP), Octavio Frias de Oliveira, ICESP, São Paulo, SP, Brazil and Plataforma Brasil, and written informed consent was obtained from all participants. The methods the present study were performed in accordance with the approved guidelines, and experimental protocols were approved by the ethics committee of the ICESP. Saliva samples were obtained voluntarily from healthy individuals ( $n = 10$ ) and oral squamous cell carcinoma (OSCC) patients ( $n = 24$ ), including patients who had undergone surgical resection (designated as “no lesion,”  $n = 10$ ) and patients who had active oral malignant lesions (designated as “lesion,”  $n = 14$ ) at the time of saliva collection. The clinical and pathological information for these oral cancer patients is summarized in Supplemental Table 1. Individuals first rinsed their mouths with 5 mL of drinking water and then harvested the saliva into a glass recipient. The saliva samples were aliquoted in 2 mL tubes and immediately frozen at  $-80^{\circ}\text{C}$ . The samples remained at  $-80^{\circ}\text{C}$  for long term storage.

**Whole saliva protein extraction.** Saliva was initially centrifuged for 5 min at  $1,500 \times g$ ,  $4^{\circ}\text{C}$  to remove intact cells and debris. A volume of  $100\ \mu\text{L}$  of whole saliva was used in the protein extraction procedure, which was performed by homogenizing the whole saliva with  $100\ \mu\text{L}$  of urea buffer (100 mM Tris-HCl pH 7.5, 8 M urea, and 2 M thiourea) containing the cOmplete Mini Protease Inhibitor Cocktail (Roche, Auckland, New Zealand), 5 mM EDTA, 1 mM PMSF and 1 mM DTT. The samples were sonicated for 10 min and centrifuged at  $10,000 \times g$  for 5 min. Total protein was quantified using a Bradford assay kit (Bio-Rad, São Paulo, Brazil).

**Salivary EV isolation and protein extraction.** Isolation of EVs was performed based on previously described methods<sup>47,48</sup>. Briefly, 1 mL of saliva was diluted at a proportion of 1:1 in phosphate buffered saline (PBS) containing protease inhibitors (1 mM PMSE, 5 mM EDTA, and cOmplete Mini Protease Inhibitor Cocktail) and 1 mM DTT and centrifuged at  $200 \times g$  for 5 min. The supernatants were centrifuged in consecutive steps of  $2,000 \times g$  for 10 min,  $3,500 \times g$  for 10 min, and  $10,000 \times g$  for 90 min, followed by a final ultracentrifugation step at  $100,000 \times g$  for 90 min. All centrifugation steps were performed at  $4^{\circ}\text{C}$ . Isolated EVs were maintained at  $-80^{\circ}\text{C}$  in PBS for microscopic and nanoparticle tracking analysis or were suspended in  $30\ \mu\text{L}$  of urea buffer (50 mM Tris-HCl pH 7.5, 8 M urea, 2 M thiourea, and 1 mM DTT) for protein extraction. Samples in urea buffer were sonicated for 5 min at room temperature. Protein quantification was performed using a Bradford assay kit (Bio-Rad, São Paulo, Brazil).

**Immunoblotting detection.** The presence of the EV marker flotilin-1 was verified by western blotting. Whole saliva and EV protein extracts were separated by 12.5% SDS-PAGE and transferred onto a nitrocellulose membrane (GE Healthcare, WI, USA), which was blocked in 1% BSA for 2 h at room temperature followed by incubation with the primary rabbit anti-flotilin-1 antibody (1:5000) (Sigma-Aldrich, MO, USA) and the secondary horseradish peroxidase-conjugated anti-rabbit IgG antibody (1:2000)

(Sigma-Aldrich, MO, USA). Protein detection was performed by chemiluminescence using an ECL kit (Amersham Biosciences, NJ, USA).

**Transmission electron microscopy.** Salivary EV preparations ( $1 \times 10^8$  particles·mL<sup>-1</sup>) were suspended in PBS and adsorbed onto glow positively charged (15 mA for 25 s) holey carbon coated copper grids (Ultrathin carbon film mesh 400 copper grid). Negative staining was performed with freshly prepared 2% aqueous uranyl acetate. EVs were imaged using a JEM-3010 transmission electron microscope equipped with a tungsten filament and operated at an acceleration voltage of 300 kV.

**Nanoparticle tracking analysis (NTA).** NTA measurements of the EVs were performed with a NanoSight NS300 (NanoSight, Amesbury, United Kingdom) equipped with a sample chamber with a 532-nm green laser. The samples were measured at room temperature in PBS for 60 s with camera gain adjustments appropriate for the correct focusing of the EVs. The software NTA 2.3 Build 0013 was used for data capture and analysis. Nanoparticle sizing and concentrations were determined based on the Brownian motion of the individual particles and light scattering measurements.

**Mass spectrometry analysis.** Whole saliva protein (10 µg) and salivary EV protein (2 µg) extracts were reduced, alkylated, trypsin digested and desalted according to previous protocols<sup>49,50</sup>. Tryptic-digested peptides were dried in a speed-vac instrument and identified in an ETD enabled LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) connected to the EASY-nLC system (Proxeon Biosystem, West Palm Beach, FL, USA) through a Proxeon nanoelectrospray ion source in a data dependent mode. Peptides were separated by a 2–30% acetonitrile gradient in 0.1% formic acid using an analytical PicoFrit Column (20 cm × ID75 µm, 5 µm particle size, New objective) at a flow rate of 300 nl/min over 40 min. The nanoelectrospray voltage was set to 2.2 kV, and the source temperature was 275 °C. All instrument methods for the LTQ Orbitrap Velos were run in the data-dependent analysis (DDA) mode. The full scan MS spectra ( $m/z$  300–1,600) were acquired in the Orbitrap analyzer after accumulation to a target value of  $1e^6$ . The resolution in the Orbitrap was set to  $r = 60,000$ . The 20 most intense peptide ions with charge states of  $\geq 2$  were sequentially isolated to a target value of 5,000 and fragmented in the linear ion trap by low-energy CID (normalized collision energy of 35%). The signal threshold for triggering an MS/MS event was set to 1,000 counts. Dynamic exclusion was enabled with an exclusion size list of 500, exclusion duration of 60 s, and a repeat count of 1. An activation  $q = 0.25$  and activation time of 10 ms were used.

**Proteomics data analysis.** Protein identification was performed with the Andromeda search algorithm within the MaxQuant v.1.3.0.3 software against the UniProt Human Protein Database (Release: June 2013, 88,820 sequences). The search parameters were set to a maximum of two missing cleavages, with a maximum error tolerance of 6 ppm for MS search and 0.5 Da for MS/MS search. Acetylation and Oxidation were set as variable modifications, and Carbamidomethylation was set as a fixed modification. Label-free protein quantification was performed using a previously described label-free quantification (LFQ) algorithm implemented in the MaxQuant software with a 2 min window for matching between runs and maximum 1% peptide and 1% protein FDR<sup>51,52</sup>. Protein intensity values were normalized using the LFQ algorithm available through the MaxQuant program and used to further identify differentially expressed proteins. Statistical analysis of the data was performed using Perseus v.1.2.7.4 software. Protein identification datasets were pre-processed for the exclusion of contaminant entries and reverse sequences identification and only identified by site entries. The proteome dataset corresponding to sample T22 was excluded from further statistical analysis due to its outlier behavior (modified z-score > 3) considering the number of proteins identified by mass spectrometry. LFQ intensity values were log<sub>2</sub> transformed; the dataset was filtered by minimum valid values (unless otherwise indicated in the text, the minimum valid values filter was set to 3) in at least one group, and statistical significance was assessed by applying a one-way ANOVA method to identify differentially expressed proteins. Storey's method<sup>53</sup> (FDR < 0.05 and bootstrap estimation of  $\lambda$ ) was further applied for q-value calculations through multiple hypothesis testing correction of calculated ANOVA p-values.

Additionally, data classification and feature selection analysis were performed using the Weka v.3 suite (<http://www.cs.waikato.ac.nz>). The support vector machine (SVM) method was applied using a sequential minimal optimization (SMO) algorithm (Polykernel -c250007 -E1.0) for training the SVM, using 66% dataset split as the training dataset, which includes the intensity values of the protein hits normalized by the z-score. The feature selection method Greedy Stepwise (forwards; -T -1.7976931348623157E308 -N -1) with Filtered Subset Evaluator as attribute estimator was applied for the z-score normalized protein intensity values for the dataset without any missing data. The dataset corresponding to sample T22 was excluded from this analysis due to its outlier behavior.

Functional annotation of the identified proteins was performed using IPA (Ingenuity® Systems, <http://www.ingenuity.com>). Significant interactions and pathways associated with datasets were determined by Fisher's exact test. Overrepresented gene ontology (GO) functional categories were identified through the ClueGO v.2.0.6 plugin, using the hypergeometric test and p-value correction with the Benjamini-Hochberg method. Global *Homo sapiens* GO annotation (GO release 12.07.2013) was used as a background annotation database, and the significance threshold was set as 0.05. Overrepresented

terms were visualized using the Cytoscape (v.2.8.3) software platform<sup>54</sup>. The raw files and protein target database were deposited in the Peptide Atlas data repository and can be accessed through the link <http://www.peptideatlas.org/PASS/PASS00317>.

**Assessment of prognostic protein candidates and functional connections associated with the identified proteins.** Protein LFQ intensity values were used to calculate the odds ratio for the proteins identified with differential abundance in the saliva of OSCC patients. The odds ratio, p-values and confidence intervals were calculated, and proteins with an odds ratio p-value of lower than 0.05 were further considered in a survival analysis. The LFQ intensity values were used for the analysis of the odds ratio for the subset of differentially expressed proteins. The median LFQ intensity values were calculated and used as a threshold for dichotomizing the values into two categories associated with low or high intensity mean values using the SPSS (Statistical Package for the Social Sciences) 17 software package. Survival curves were created using the Kaplan-Meier method<sup>55</sup> with comparisons by the log-rank test. Furthermore, our analysis of possible connections between the differentially expressed proteins and small bioactive molecules was performed using the connectivity map (cMAP; <https://www.broadinstitute.org/cmap/>) build 2 database. Salivary proteins that were up- and downregulated in OSCC patients were used as input data for the prediction of common gene expression changes that may occur in response to drugs, genes or proteins. Several drugs were predicted to influence the expression of these candidate protein coding genes.

Additionally, we performed a cross-cancer analysis of the frequency of genetic alterations (mutation, deletion, amplification and multiple alterations) in the genes that code for the differentially expressed proteins in the saliva of OSCC patients using cBioPortal ([www.cbioportal.org](http://www.cbioportal.org)). Moreover, protein-coding genes that corresponded to differentially expressed proteins identified through our proteomic approach and the proteins with potential for data classification were used as input data for the prediction of their roles as oncogenes or tumor suppressor genes in different types of cancer using the S-score system ([www.bioinformatics-brazil.org/S-score/](http://www.bioinformatics-brazil.org/S-score/))<sup>56</sup>.

## References

- van der Waal, I. Are we able to reduce the mortality and morbidity of oral cancer; some considerations. *Med Oral Patol Oral Cir Bucal* **18**, e33–37, doi: 10.4317/medoral.18486 (2013).
- Hauw, J. J., Boutry, J. M., Guillermin, A. M., Barbe, M. P. & Boissonnet, H. A new promotor of nerve growth: naftidrofuryl. *C R Acad Sci III* **302**, 543–548 (1986).
- Lippman, S. M., Sudbo, J. & Hong, W. K. Oral cancer prevention and the evolution of molecular-targeted drug development. *J Clin Oncol* **23**, 346–356, doi: 10.1200/JCO.2005.09.128 (2005).
- Sudbo, J. Novel management of oral cancer: a paradigm of predictive oncology. *Clin Med Res* **2**, 233–242 (2004).
- Carvalho, A. L., Nishimoto, I. N., Califano, J. A. & Kowalski, L. P. Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database. *Int J Cancer* **114**, 806–816, doi: 10.1002/ijc.20740 (2005).
- Jerjes, W. *et al.* Clinicopathological parameters, recurrence, locoregional and distant metastasis in 115 T1-T2 oral squamous cell carcinoma patients. *Head Neck Oncol* **2**, 9, doi: 10.1186/1758-3284-2-9 (2010).
- da Silva, S. D. *et al.* Advances and applications of oral cancer basic research. *Oral Oncol* **47**, 783–791, doi: 10.1016/j.oraloncology.2011.07.004 (2011).
- Yan, W. *et al.* Systematic comparison of the human saliva and plasma proteomes. *Proteomics Clin Appl* **3**, 116–134, doi: 10.1002/prca.200800140 (2009).
- Cheng, L., Sharples, R. A., Scicluna, B. J. & Hill, A. F. Exosomes provide a protective and enriched source of miRNA for biomarker profiling compared to intracellular and cell-free blood. *J Extracell Vesicles* **3**, doi: 10.3402/jev.v3.23743 (2014).
- Markopoulos, A. K., Michailidou, E. Z. & Tzimagiorgis, G. Salivary markers for oral cancer detection. *Open Dent J* **4**, 172–178, doi: 10.2174/1874210601004010172 (2010).
- D'Souza-Schorey, C. & Clancy, J. W. Tumor-derived microvesicles: shedding light on novel microenvironment modulators and prospective cancer biomarkers. *Genes Dev* **26**, 1287–1299, doi: 10.1101/gad.192351.112 (2012).
- Choi, D. S., Kim, D. K., Kim, Y. K. & Cho, Y. S. Proteomics, transcriptomics and lipidomics of exosomes and ectosomes. *Proteomics* **13**, 1554–1571, doi: 10.1002/pmic.201200329 (2013).
- Tang, H., Wu, Z., Zhang, J. & Su, B. Salivary lncRNA as a potential marker for oral squamous cell carcinoma diagnosis. *Mol Med Rep* **7**, 761–766, doi: 10.3892/mmr.2012.1254 (2013).
- Shpitzer, T. *et al.* Salivary analysis of oral cancer biomarkers. *British journal of cancer* **101**, 1194–1198, doi: 10.1038/sj.bjc.6605290 (2009).
- Jessie, K. *et al.* Aberrant proteins in the saliva of patients with oral squamous cell carcinoma. *Electrophoresis* **34**, 2495–2502, doi: 10.1002/elps.201300107 (2013).
- Sivadasan, P. *et al.* Human salivary proteome - a resource of potential biomarkers for oral cancer. *Journal of proteomics*, doi: 10.1016/j.jprot.2015.05.039 (2015).
- Arellano-Garcia, M. E. *et al.* Multiplexed immunobead-based assay for detection of oral cancer protein biomarkers in saliva. *Oral Dis* **14**, 705–712, doi: 10.1111/j.1601-0825.2008.01488.x (2008).
- Leung, E. L. *et al.* Non-small cell lung cancer cells expressing CD44 are enriched for stem cell-like properties. *PLoS One* **5**, e14062, doi: 10.1371/journal.pone.0014062 (2010).
- Yanagawa, H. *et al.* Serum levels of interleukin 6 in patients with lung cancer. *Br J Cancer* **71**, 1095–1098 (1995).
- Zhang, Y. *et al.* Potential mechanism of interleukin-8 production from lung cancer cells: an involvement of EGF-EGFR-PI3K-Akt-Erk pathway. *J Cell Physiol* **227**, 35–43, doi: 10.1002/jcp.22722 (2012).
- Lamb, J. The Connectivity Map: a new tool for biomedical research. *Nat Rev Cancer* **7**, 54–60, doi: 10.1038/nrc2044 (2007).
- Cecilian, F., Giordano, A. & Spagnolo, V. The systemic reaction during inflammation: the acute-phase proteins. *Protein Pept Lett* **9**, 211–223 (2002).
- Dowling, P. *et al.* Analysis of acute-phase proteins, AHSG, C3, CLI, HP and SAA, reveals distinctive expression patterns associated with breast, colorectal and lung cancer. *International journal of cancer. Journal international du cancer* **131**, 911–923, doi: 10.1002/ijc.26462 (2012).

24. Denny, P. *et al.* The proteomes of human parotid and submandibular/sublingual gland salivas collected as the ductal secretions. *Journal of Proteome Research* **7**, 1994–2006. doi: 10.1021/pr700764j (2008).
25. Lengyel, E. Ovarian cancer development and metastasis. *Am J Pathol* **177**, 1053–1064. doi: 10.2353/ajpath.2010.100105 (2010).
26. Cho, M. S. *et al.* Autocrine effects of tumor-derived complement. *Cell Rep* **6**, 1085–1095. doi: 10.1016/j.celrep.2014.02.014 (2014).
27. Lim, J. *et al.* C5aR and C3aR antagonists each inhibit diet-induced obesity, metabolic dysfunction, and adipocyte and macrophage signaling. *FASEB J* **27**, 822–831. doi: 10.1096/fj.12-220582 (2013).
28. Kindt, N. *et al.* Macrophage migration inhibitory factor in head and neck squamous cell carcinoma: clinical and experimental studies. *J Cancer Res Clin Oncol* **139**, 727–737. doi: 10.1007/s00432-013-1375-7 (2013).
29. Kindt, N. *et al.* Pharmacological inhibition of macrophage migration inhibitory factor interferes with the proliferation and invasiveness of squamous carcinoma cells. *Int J Oncol* **43**, 185–193. doi: 10.3892/ijo.2013.1944 (2013).
30. Carli, C., Metz, C. N., Al-Abed, Y., Naccache, P. H. & Akoum, A. Up-regulation of cyclooxygenase-2 expression and prostaglandin E2 production in human endometriotic cells by macrophage migration inhibitory factor: involvement of novel kinase signaling pathways. *Endocrinology* **150**, 3128–3137. doi: 10.1210/en.2008-1088 (2009).
31. Calandra, T. & Roger, T. Macrophage migration inhibitory factor: a regulator of innate immunity. *Nat Rev Immunol* **3**, 791–800. doi: 10.1038/nri1200 (2003).
32. Rakoff-Nahoum, S. Why cancer and inflammation? *Yale J Biol Med* **79**, 123–130 (2006).
33. Peinado, H. *et al.* Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat Med* **18**, 883–891. doi: 10.1038/nm.2753 (2012).
34. Shao, H. *et al.* Protein typing of circulating microvesicles allows real-time monitoring of glioblastoma therapy. *Nat Med* **18**, 1835–1840. doi: 10.1038/nm.2994 (2012).
35. Rokavec, M. *et al.* IL-6R/STAT3/miR-34a feedback loop promotes EMT-mediated colorectal cancer invasion and metastasis. *J Clin Invest* **124**, 1853–1867. doi: 10.1172/JCI73531 (2014).
36. Li, J., Ji, Z., Qiao, C., Qi, Y. & Shi, W. Overexpression of ADAM9 Promotes Colon Cancer Cells Invasion. *J Invest Surg*. doi: 10.3109/08941939.2012.728682 (2013).
37. Micocci, K. C. *et al.* ADAM9 silencing inhibits breast tumor cell invasion *in vitro*. *Biochimie* **95**, 1371–1378. doi: 10.1016/j.biochi.2013.03.001 (2013).
38. Salama, I., Malone, P. S., Mihaimed, F. & Jones, J. L. A review of the S100 proteins in cancer. *Eur J Surg Oncol* **34**, 357–364 (2008).
39. Choi, D. K. *et al.* Clinicopathological roles of S100A8 and S100A9 in cutaneous squamous cell carcinoma *in vivo* and *in vitro*. *Arch Dermatol Res*. doi: 10.1007/s00403-014-1453-y (2014).
40. Namura, T. *et al.* [Possible mechanism for regulation of inflammatory responses with the S100A8/A9 protein]. *Rinsho Byori* **58**, 651–657 (2010).
41. Nigro, P., Pompilio, G. & Capogrossi, M. C. Cyclophilin A: a key player for human disease. *Cell Death Dis* **4**, e888. doi: 10.1038/cddis.2013.410 (2013).
42. Pfeffer, S. R., Dirac-Svejstrup, A. B. & Soldati, T. Rab GDP dissociation inhibitor: putting rab GTPases in the right place. *J Biol Chem* **270**, 17057–17059 (1995).
43. Casalou, C. *et al.* Arl13b and the non-muscle myosin heavy chain IIA are required for circular dorsal ruffle formation and cell migration. *J Cell Sci*. doi: 10.1242/jcs.143446 (2014).
44. Morin, N. A. *et al.* Nonmuscle myosin heavy chain IIA mediates integrin LFA-1 de-adhesion during T lymphocyte migration. *J Exp Med* **205**, 195–205. doi: 10.1084/jem.20071543 (2008).
45. Jacobelli, J., Estin Matthews, M., Chen, S. & Krummel, M. F. Activated T cell trans-endothelial migration relies on myosin-IIA contractility for squeezing the cell nucleus through endothelial cell barriers. *PLoS One* **8**, e75151. doi: 10.1371/journal.pone.0075151 (2013).
46. Derycke, L. *et al.* The role of non-muscle myosin IIA in aggregation and invasion of human MCF-7 breast cancer cells. *Int J Dev Biol* **55**, 835–840. doi: 10.1387/ijdb.1133361d (2011).
47. Xiao, H. & Wong, D. T. Proteomic analysis of microvesicles in human saliva by gel electrophoresis with liquid chromatography-mass spectrometry. *Anal Chim Acta* **723**, 61–67. doi: 10.1016/j.aca.2012.02.018 (2012).
48. Gallo, A., Tandon, M., Alevizos, I. & Illei, G. G. The majority of microRNAs detectable in serum and saliva is concentrated in exosomes. *PLoS One* **7**, e30679. doi: 10.1371/journal.pone.0030679 (2012).
49. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nano-electrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* **75**, 663–670 (2003).
50. Villen, J. & Gygi, S. P. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat Protoc* **3**, 1630–1638. doi: 10.1038/nprot.2008.150 (2008).
51. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367–1372. doi: 10.1038/nbt.1511 (2008).
52. Lubner, C. A. *et al.* Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. *Immunity* **32**, 279–289. doi: 10.1016/j.immuni.2010.01.013 (2010).
53. Storey, J. D., Taylor, J. E. & Siegmund, D. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *Journal of the Royal Statistical Society B* **66**, 187–205 (2004).
54. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432. doi: 10.1093/bioinformatics/btq675 (2011).
55. Goel, M. K., Khanna, P. & Kishore, J. Understanding survival analysis: Kaplan-Meier estimate. *Int J Ayurveda Res* **1**, 274–278. doi: 10.4103/0974-7788.76794 (2010).
56. de Souza, J. E. *et al.* S-score: a scoring system for the identification and prioritization of predicted cancer genes. *PLoS One* **9**, e94147. doi: 10.1371/journal.pone.0094147 (2014).

## Acknowledgements

This work was supported by FAPESP Grants 2009/54067-3, 2010/19278-0, and CNPq Grants 470549/2011-4, 301702/2011-0 and 470268/2013-1. We thank all donors for kindly providing biological samples to this study. We thank Dr. Rodrigo Portugal and Dr. Alexandre Cassago from the Laboratório Nacional de Nanotecnologia (LNNano), Campinas, Brazil, for their support on the transmission electron microscopy experiments.

## Author Contributions

Draft the manuscript: F.V.W. and A.F.P.L.; Performed the sample preparation, experiments and electron microscopy: F.V.W.; Performed the LC-MS/MS: R.R.D. and F.V.W.; Performed the proteomics data analysis: F.V.W., L.L., D.M.R.P., A.F.P.L. and C.R.; Performed the clinical sample collection: A.C.P.R.,

R.D.C., A.R.S. and T.B.B.; Performed anatomopathological classification and gathered clinical information: A.C.P.R., R.D.C., A.R.S., T.B.B. and A.F.G. All authors have read and agreed with the final version of the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Winck, F. V. *et al.* Insights into immune responses in oral cancer through proteomic analysis of saliva and salivary extracellular vesicles. *Sci. Rep.* **5**, 16305; doi: 10.1038/srep16305 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>