Current advances in identification of cancer biomarkers in saliva

Martha Arellano1, Jiang Jiang1, Xiaofeng Zhou2, Lifeng Zhang1, Hui Ye2, David T. Wong1,3, Shen Hu1,3

1School of Dentistry and Dental Research Institute, University of California at Los Angeles, Los Angeles CA 90095, 2Center for Molecular Biology of Oral Diseases, College of Dentistry, University of Illinois at Chicago, Chicago IL 60612, 3Jonsson Comprehensive Cancer Center, University of California at Los Angeles, Los Angeles CA 90095

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Genomic targets in saliva
   3.1. DNA
      3.1.1. Mutation
      3.1.2. DNA methylation
      3.1.3. Viral DNA
   3.2. RNA
4. Proteomic targets in saliva
   4.1. Systemic identification in salivary proteins
   4.2. Elevated levels of specific protein molecules
   4.3. Antibodies to mutant proteins and viral targets
5. Perspective
6. Acknowledgements
7. References

1. ABSTRACT

Saliva samples from certain disease patients harbor a wide spectrum of proteins, mRNAs, DNAs and other molecules that may be associated with the disease phenotype. If successfully discovered and validated, these informative molecules may serve as biomarkers, leading to the use of non-invasive biofluid for detecting and monitoring the diseases. This article summarizes the current advances in searching for potential biomarkers in saliva for human cancers, especially head and neck/oral cancers. With the new molecular profiling technologies such as microarray and proteomics, we are expecting to reveal highly discriminatory genomic and proteomic targets that can best detect the disease status.

2. INTRODUCTION

Biomarker refers to measurable and quantifiable biological parameters that can serve as indicators for health and physiology-related assessments, such as pathogenic processes, environmental exposure, disease diagnosis and prognosis, or pharmacologic responses to a therapeutic intervention. If successfully developed, these informative parameters may lead to clinical tools for early detection of diseases, surveillance of disease recurrence and predicting disease aggressiveness. Biomarkers can also provide vital clues regarding drug effectiveness and toxicity or classify patient subgroups to facilitate the development of molecular targeted therapy for specific patients.
One might think biopsies of disease tissues are the ideal sample source for biomarker discovery research. However, in terms of diagnostic purpose, body fluids (e.g., saliva and blood) are more attractive because of simple and minimally invasive sample collection. For patients, the non-invasive collection procedure for saliva dramatically reduces anxiety and discomfort and simplifies procurement of repeated samples for monitoring over time. Compared to tissue biopsies, saliva is an easily accessible fluid and, therefore, a large number of saliva samples can be enrolled for clinical studies. This allows enough statistical power for a robust study design, and true diagnostic signatures can be revealed as disease biomarkers.

Saliva harbors a wide spectrum of analytes (proteins, mRNAs, DNAs and other small molecules) that may reflect the health/disease status in human body. Profiling of these molecules in saliva from a disease population can potentially yield robust and reproducible clinical parameters for diagnosis and prognosis of the disease. Advances in microarray and proteomic technologies have enabled large-scale profiling of gene expression in cells, tissue and body fluids. When applied to clinical studies towards biomarker discovery, such a global analysis allows attaining the most discriminatory clinical parameters for diagnosis and prognosis of the disease. Advances in microarray and proteomic technologies have enabled large-scale profiling of gene expression in cells, tissue and body fluids. When applied to clinical studies towards biomarker discovery, such a global analysis allows attaining the most discriminatory biomolecules that can best detect the disease status. In this article, we aim to summarize the current efforts on identification of saliva genomic and proteomic targets as biomarkers for human cancers, particularly in head and neck/oral cancer. Knowing the current status will help advance future investigations in this emerging field.

3. GENOMIC TARGETS IN SALIVA

3.1. DNA

3.1.1. Mutation
Cancer is caused by the accumulation of genetic or epigenetic changes that lead to the activation of oncogenes and inactivation of tumor suppressor genes. Monitoring these changes in saliva (or saliva rinse) from patients with cancers, especially head and neck cancer, represents a promising approach to discover specific markers for the cancer detection (1, 2). For instance, tumor-specific p53 mutations were identified in 5 out of 7 (71%) saliva samples from patients with head and neck squamous cell carcinoma (HNSCC). For every 100 cells in the saliva, two or three cells had exfoliated from the patient's cancer and contained the same p53 mutation (1). Microsatellite analysis of tumor-specific genetic alterations in saliva and swab samples was also investigated for the purpose of detecting and monitoring HNSCC (3). In this study, microsatellite instability was detectable in the saliva in 24 (96%) of 25 cases with known cancer, and loss of heterozygosity was identified in the test sample in 19 (61%) of 31 cases. On the other hand, no microsatellite alterations appeared in any of the samples from the healthy control subjects.

The mutated mitochondrial DNA (mtDNA) has also been useful targets for the detection of HNSCC in saliva. The mutations of mtDNA were identified in 46% of head and neck cancers, and importantly, the same mtDNA mutations were detected in 67% (6 of 9) of saliva samples from head and neck cancer patients (4). Quantitative PCR analysis of cytochrome c oxidase I (Cox I) and cytochrome c oxidase II (Cox II) genes in oral rinse samples suggested mtDNA content alteration was associated with HNSCC independently of age and smoking exposure, and advanced stage of HNSCCs exhibited higher level of mtDNA content in saliva. Similar approach was also used to measure changes in mtDNA content in pretreatment and posttreatment saliva rinse samples obtained from patients undergoing surgical resection for primary HNSCC. The study suggested that salivary mtDNA content was decreased in never smokers and in response to radiation therapy after primary surgical resection (5, 6).

3.1.2. DNA methylation
Aberrant promoter hypermethylation is common in head and neck cancer and may provide another potential target for detecting exfoliated cancer cells in saliva. By using methylation-specific PCR to search for promoter hypermethylation of the tumor suppressor gene p16, the DNA repair gene O6-methylguanine-DNA-methyltransferase (MGMT) and the putative metastasis suppressor gene death-associated protein kinase (DAPK), Rosas et al. identified aberrant methylation of at least one of these genes in 17 (56%) of 30 head and neck primary tumors; 14 (47%) of 30 at p16, 10 (33%) of 30 at DAPK and 7 (23%) of 30 at MGMT. However, Of 30 saliva samples from healthy control subjects (15 smokers and 15 nonsmokers), only one sample from a smoking patient was positive for DNA methylation at two target genes (7). In a recent study to examine the methylation of six genes (TIMP3, ECAD, p16, MGMT, DAPK, and RASSF1) in saliva, similar methylation profiles were observed in newly diagnosed and second primary cancers. Within the 22 patients followed after treatment, abnormal methylation was detectable in the saliva of 5 patients few months before the clinical signs of relapse, allowing curable surgery. For the other 17 patients whose saliva samples were negative, 16 were in remission and only 1 relapsed. These results suggested that gene methylation in saliva is potentially useful for the follow-up and early detection of relapses in HNSCC patients (8).

3.1.3. Viral DNA
The Epstein-Barr virus (EBV) is a common human virus that causes infectious mononucleosis and plays a role in the emergence of two forms of cancer: Burkitt's lymphoma, and nasopharyngeal carcinoma (NPC). The levels of EBV DNA in saliva can be readily detectable using PCR assays, allowing a simple method to diagnose the viral infection and correlate the prevalence of EBV in saliva with tissue biopsies (9-11). Human papilloma virus (HPV) 16 is likely to be an etiologic factor in a subset of HNSCC. A quantitative PCR-based assay was investigated to detect HPV 16 in salivary rinses as a screening method for HNSCC. A total of 45.6% (42/92) of primary HNSCC and 32.6% (30/92) of saliva rinse samples from HNSCC patients had any detectable HPV 16 DNA. HPV16 was detected in 50% (21/42) of the saliva rinse samples from HNSCC patients with detectable HPV16 level in their tissue, 18% (9/50) of saliva rinse samples from patients with HPV 16 negative primary HNSCC, in 2.8% (17/604)
Cancer biomarkers in saliva

of the normal controls. Nonsmokers had significant higher HPV 16 level than smokers in the cohort of cancer patients. These results suggested that HPV 16 DNA in saliva rinses can reflect HPV 16 status of primary HNSCC (12). Human herpesvirus 8 (HHV-8) has been proposed as a sexually transmitted etiologic agent of Kaposi's sarcoma (KS). Similarly, PCR assay has been used to detect HHV-8 DNA in saliva (37% prevalence) from KS patients. The average number of HHV-8 copies per microgram of positive target DNA was 33,000 for saliva, as compared to 64, 000, 40, and 300 for skin, plasma, and semen samples, respectively (13).

3.2. RNA

We discovered that human saliva contains cell-free mRNAs (salivary transcriptome) that can be highly informative and discriminatory for disease detection. By using Affymetrix U133 microarrays for genome-wide profiling of saliva from 10 healthy subjects, we found all subjects have approximately 3,000 different mRNA molecules in their saliva. An average of 3,143 probe sets on each array was assigned, representing approximately 3,000 different mRNAs per individual (14). We further investigated the source, integrity and stability of saliva mRNAs, and found they originate from both nucleus and mitochondria (15). In order to optimize sample processing or stabilization, we compared the stability of RNA from whole saliva at room temperature with three stabilizing reagents, Superase Inhibitor (Invitrogen), RNAlater (Ambion) or the RNAprotect Saliva reagent (RPS, Qiagen). The investigation clearly showed that overall RPS provides the best preservation of salivary mRNAs (16).

The discovery of salivary transcriptome, coupled with the ability to harness genomic information by high-throughput technology platforms such as DNA microarrays, ideally positions salivary genomic targets for applications in human disease detection. This was demonstrated in our study on discovery of salivary mRNA markers for detection of oral squamous cell carcinoma (OSCC). Initial microarray profiling followed by real-time quantitative PCR revealed that four salivary mRNAs (OAZ, SAT, IL8, and IL1b) collectively have a discriminatory power of 91% sensitivity and specificity for oral cancer detection, and the area under the ROC curve measuring 95% (17). The clinical value of these salivary mRNA markers for oral cancer detection was also compared with the serum-based detection developed with the same methodology and patient cohorts (18). Use the same patient cohort, the panel of four salivary mRNA biomarkers collectively exhibited a ROC value of 0.95 whereas a panel of serum-based RNA biomarkers exhibited a ROC value of 0.88. These results show that saliva testing demonstrated a slightly better diagnostic value than blood for oral cancer detection.

4. PROTEOMIC TARGETS IN SALIVA

4.1. Systemic identification in salivary proteins

Whole saliva of a healthy adult typically has about 1 mg/ml of total proteins, which are expressed and secreted from major and minor glands. These polypeptides not only play important roles in maintaining oral and general health but may also serve as biomarkers to monitor normal health and disease status. An in-depth characterization of the human saliva proteome (the total proteins/peptides present in saliva) can therefore provide a valuable resource for researchers within the fields of oral biology and saliva diagnostics. To date, we have identified more than 1000 proteins in human whole saliva by using the combination of a variety of prefractionation techniques with quadrupole time-of-flight (QTOF) mass spectrometry (MS) (19-21). Figure 1 illustrates the identification of salivary proteins using either 2-D gel electrophoresis (2-DE)/MS or shotgun proteomics based on LC-MS/MS. Other research groups have used similar approaches to fractionate proteolytic peptides from whole saliva proteins based on the coupling of capillary isoelectric focusing or free-flow electrophoresis with LC-MS/MS to decipher the saliva proteome (22, 23); again, approximately 1000 proteins have been catalogued. With a 4-year collaborative effort from a NIDCR-funded consortium of three laboratories from Scripps, UCSF and UCLA to decipher human saliva proteome, more than 1100 non-redundant proteins were identified from major salivary glandular secretions, including parotid and submandibular/sublingual saliva (24). Whole saliva glycoproteins have also been identified based on the use of hydrazide chemistry and LC-MS/MS (25). A central repository has been established (www.hspp.ucla.edu) to consolidate the acquired proteomic data and link the identified proteins to public annotated protein databases.

4.2. Elevated levels of specific protein molecules

Previous studies have revealed a panel of saliva proteins from patients with HNSCC/OSCC may have diagnostic potential for the diseases. Salivary soluble CD44 (sCD44) and hyaluronidase (HAase) were both found at elevated levels in the patients with HNSCCs (26, 27). The sCD44 levels could be detected in 79% of mucosally invasive HNSCC, and did not vary significantly with tumor size, stage, recurrence, history of radiation treatment, or tobacco and alcohol risk factors (26). Hyaluronic acid (HA), a component of extracellular matrix, has been suggested to play a role in the metastasis of HNSCC. HA fragments are generated when hyaluronidase (HAase), an endoglycosidase, degrades the HA polymer. The levels of HA and HAase in HNSCC patients were found to be 4.9-fold and 3.7-fold higher than normal controls (27). The telomerase activity, as measured by telomerase PCR-ELISA method, was positively detected in 75.0% of patients with OSCC and only in 6.67% of normal individuals. However, no significant difference in telomerase activity was observed between the early and late stage OSCC patients, nor between patients with and without lymph nodes metastasis (28). Nagler et al examined the salivary levels of epithelial serum tumor markers: cytokeratin 19 fragment Cyfra21-1, tissue polypeptide antigen (TPA), and CA125. These molecules were found at ~ 4 fold increases in OSCCs over the controls. Strategically combination of these markers resulted in sensitivity, specificity, and negative and positive predictive values of 71%, 75%, 71% and 75%, respectively. These values are similar to those obtained when measuring
Cancer biomarkers in saliva

Figure 1. Analysis and identification of saliva proteins using mass spectrometry-based proteomics. (A) 2-D gel electropherogram of proteins in whole saliva. The circled gel spots represent the glycoforms of carbonic anhydrase VI. These gel spots were excised and identification of the proteins was performed using in-gel digestion and LC-MS/MS. (B) LC-ESI tandem MS spectrum of a tryptic peptide, LENSLLDHR (m/z, 548.73, 2+), originated from carbonic anhydrase VI. (C) Base-peak chromatogram of the total peptides resulting from direct tryptic digestion of whole saliva proteins. (D) LC-MS/MS spectrum of a tryptic peptide, FGQGSPIVLDDVR (m/z, 730.40, 2+), from DMBT1 (Deleted in Malignant Brain Tumors 1, a.k.a., Glycoprotein 340).

Defensins are a family of defensive peptides with antimicrobial and cytocidal properties. Human alpha-defensin 1 (HNP-1) level was found significantly higher in OSCC patients than normal controls, however it dropped in OSCC patients by about 50% after surgery (30). Guanylin is a peptide synthesized by the duct epithelial cells in the salivary glands. It is highly expressed in the human parotid and submandibular glands and in most relevant tumors. Guanylin was detected in saliva at a considerably higher than in circulation system (31). Using MS-based proteomics approaches, we recently discovered potential protein biomarkers from whole saliva of OSCC patients, and subsequent validation by immunoassays suggested these protein biomarkers are valid for OSCC detection (32, 33). Using SDS-PAGE and LC-MS/MS, Ohshiro et al. found that alpha-1-B-glycoprotein and complement factor B proteins were only detected in patients with HNSCC while cystatin S, parotid secretory factor, and poly-4-hydrolase beta-subunit proteins were detected in most normal saliva but not in patient specimens (34).

Several studies also indicated that salivary NF-kappaB-dependent cytokines, TNF-alpha, interleukin (IL) 6, IL1-beta and IL8, were elevated in OSCC patients as compared to healthy controls (35-37). Increased levels of these cytokines were also observed in patients with oral premalignant lesions (OPML) or oral lichen planus patients as compared to controls (37-39). However, salivary levels of IL-6 and TNF-alpha were not influenced by smoking habits, and did not correlate with the size of leukoplakia (lesions), or with its anatomic localization (39). Nevertheless, whether these cytokines can be used as cancer biomarkers remains to be a question. Their clinical utility for OSCC may need to be further confirmed by including an inflammatory disease group (e.g., periodontal disease) in study design to ensure their expression alteration is tightly associated with the cancer.

Breast cancer is the second deadly cancer for the U.S. women and the early diagnosis is essential for successful treatment of patients with the disease. An early study by Navarro et al reported that epidermal growth factor (EGF) levels in saliva samples from breast cancer patients were significantly higher than healthy women (40). However, EGF levels remain elevated even after successfully treatment and remission of the disease. This finding somehow undermines the potential of using saliva EGF as breast cancer biomarker. Recent studies were focused on C-erbB-2, (Her-2/neu), and cancer antigen CA15-3, which are found present at higher levels in whole saliva of patients with breast cancer than in controls and therefore may have clinical implications (41-44). C-erbB-2, a 185-kD transmembrane phosphoglycoprotein, is a carcinoma marker in tissue biopsies diagnosed as malignant breast cancer, while CA 15-3 is a mucinous glycoprotein, usually used as a "gold standard" against which to compare the c-erbB-2 protein's diagnostic effectiveness. Although
Cancer biomarkers in saliva

the mechanism for such larger protein like c-erbB-2 to enter saliva is not clear, the elevated levels of these biomarkers have been reported in patients with malignant cancer compared with healthy controls and benign subjects (41). It was also reported that the levels of the cancer biomarkers CA 15-3 and c-erbB-2 are not affected by tobacco usage, menopausal status, estrogen usage, systemic diseases, prescription medications, race, gender, age nor sampling during various times of the day (42, 43). Nevertheless, Bigler et al reported there was a significant difference in salivary c-erbB-2 level before and after therapy (44). These results suggest that c-erbB-2 level in saliva is a suitable marker for measuring patient response to chemotherapy and/or surgical treatment of their disease. In clinical practice, ~ 20% of breast cancer patients identified by this biomarker are chosen for the trastuzumab treatment, which is a humanized monoclonal antibody targeted against c-erbB-2 protein (45). If constantly validatable, salivary c-erbB-2 may be used to define which patients are most likely to benefit from such a molecular-targeted therapy.

4.3. Antibodies to mutant proteins and viral targets

Many malignant tumors are known to have mutations and other abnormalities of the p53 gene. Since mutant p53 has prolonged half-life, leading to the accumulation of p53 protein in the nuclei of the tumor cells, which in turn may result in the production of p53 antibodies directed against this protein. These antibodies can be detected in sera of patients with different types of malignancies and in the saliva of OSCC patients, which can thus assist in the early detection and screening of these tumors(46). Based on the analysis of p53 antibodies in tissue, sera and saliva of OSCC patients using immunoassays, Tavassoli et al. demonstrated that p53 antibodies against the mutant p53 proteins can be a potential marker for a subset of OSCC and other cancers with p53 aberrations (47). In addition, secretory IgA antibodies specific to EBV has been found predominant in saliva of NPC patients, suggesting their potential role as biomarkers for EBV-associated NPC (48, 49). The EBV-encoded nuclear antigen 1 (EBNA-1) is the only latent EBV antigen consistently expressed in malignant tissues of the nasopharynx. Salivary IgA antibodies against EBV peptide p107, the major epitope of EBNA-1, were detected in 39/46 (84.8%) NPC patients whereas only 3/42 (7.1%) of normal saliva samples were IgA-p107 positive (50). The EBV glycoprotein gp340 is the dominant component of the EBV envelope and a major target for the virus-neutralizing antibody response. Salivary IgA to gp340 were detected in a minority (12-19%) of healthy virus carriers whereas in a higher proportion (49%) of NPC patients (51). Lastly, the IgA antibodies to HPV-16 virus-like particles (VLP-16) were detected in saliva from 54.3% (44/81) of women with cervical neoplasia, compared with 8% (3/36) in controls. Although oral testing of IgA antibody to VLP-16 proved no more sensitive than serum antibody detection for the determination of HPV infection but could be useful as a non-invasive screening method for women with cervical neoplasia and for estimating the mucosal antibody response to HPV vaccines (52).

5. PERSPECTIVE

Salivary genomics and proteomics are clearly promising approaches for the identification of biomarkers for HNSCC/OSCC. Some interesting targets have been reported as possible biomarkers for detection and monitoring of the diseases. However, several issues exist that could complicate the translation of these salivary biomarkers into useful diagnostic and prognostic tools. First, whole saliva, in contrast to serum, is not constant in its composition because of changes in salivary flow and a differential contribution from the different salivary glands or other sources within the oral cavity (53, 54). Therefore how to unmask minor changes in certain disease-relevant markers remains to be a major challenge. We need to systematically study all the variables affecting the salivary content in human saliva and standardize the sample collection and handling procedures for saliva biomarker study. Secondly, the complexity of human saliva complicates cancer biomarker discovery. Unbiased approach is required to profile saliva components comprehensively so that the most discriminatory molecules can be unveiled as disease biomarker. For instance, human saliva contains a large number of proteins differing by an extraordinary wide concentration range. How to deplete the highly abundant proteins such as amylase and proline-rich proteins, prior to the profiling low abundance proteins, needs to be addressed. Thirdly, head and neck cancers have been the focus of most saliva-based biomarker research. Whether we can extend the applications to other major cancers depends on our understanding how the biomarker molecules enter into saliva. The biology behind the discovered targets should be well studied to bridge the molecules with the disease mechanism.

The improvement in early detection of cancer (especially OSCC) is essential for successful patient management. Considering that approximately 10% of the general population have oral mucosal abnormalities and precancerous and early cancerous lesions rarely demonstrate distinct clinical characteristics, there is a growing realization that some premalignant and early cancerous lesions are not readily detectable by visual inspection (55). Therefore, the integration of biomarker-based screening with a conventional oral examination is extremely important for early diagnosis. This obviously requires comparative genomic and proteomic analyses between patients with oral precancer and cancer in order to achieve discovery of robust markers for early detection. There is also concern about using inflammation-associated genes (e.g., cytokines) as cancer biomarkers. Perhaps a rational approach, when designing a biomarker study, is to include a second inflammatory disease group, in addition to a healthy control group, to reveal the genuine cancer-associated alterations. For instance, periodontal disease patients can be enrolled as a second control group in order to test the diagnostic value of those inflammatory proteins in OSCC.

It is also challenging to translate markers from proteomic or genomic investigations into real-world diagnostic or prognostic applications. Approval of using a
Cancer biomarkers in saliva

marker or set of markers for a given clinical decision relies on the results of large-scale multi-centric clinical trials (56), and approval of the use of the detection technology for that purpose. The successful completion of biomarker development needs to follow the guidelines set forth by EDRN, an initiative of the National Cancer Institute (NCI), which was established in response to the recent development of emerging technologies (e.g., proteomics and microarrays) for cancer screening (57). The appropriate usage of these biomarkers in clinics also relies on novel diagnostic devices, e.g., microfluidics-based chips, for simple and high-throughput measurement of these biomarkers in patients’ saliva at the point of care (58, 59). If successfully discovered and consistently validated, these salivary biomarkers coupled with microfluidic devices will become a powerful tool for screening and monitoring patients with cancers in the future.

6. ACKNOWLEDGEMENTS

This work was supported by the PHS grants R03 DE017144, R21CA122806 (to S. Hu), K22 DE014847, R03 DE016569, R03 CA114688 (to X. Zhou) and U01 DE016275 (to D. Wong).

7. REFERENCES


Cancer biomarkers in saliva


39. V. Brailo, V. Vucićević-Boras, A. Cekić-Arambas, I. Z. Alajbeg, A. Milenović and J. Lukac: The significance of
Cancer biomarkers in saliva


48. H. C. Ho, M. H. Ng, and H. C. Kwan: IgA antibodies to Epstein-Barr viral capsid antigens in saliva of nasopharyngeal carcinoma patients. **Br J Cancer**, 35, 888-890 (1977)

49. C. Desgranges, J. Y. Li and De-Thé . EBV specific secretory IgA in saliva of NPC patients. Presence of secretory piece in epithelial malignant cells. **Int J Cancer**, 20, 881-886 (1977)


**Key Words**: Cancer biomarker, Saliva Proteome, Saliva Transcriptome, Microarray, Proteomics, Review

**Send correspondence to**: Shen Hu, CHS 63-070, 10833 Le Conte Ave, Los Angeles, CA 90095-1668, Tel: 310-206-8834, Fax: 310-794-7109, E-mail: shenhu@ucla.edu

http://www.bioscience.org/current/vol1S.htm