Critical Review of Literature
Stomatology

Diagnostic potential of saliva proteome analysis: a review and guide to clinical practice

Camilla Vieira ESTEVES[a]
Wladimir Gushiken de CAMPOS[a]
Milena Monteiro de SOUZA[b]
Silvia Vanessa LOURENÇO[c]
Walter Luiz SIQUEIRA[d]
Celso Augusto LEMOS-JÚNIOR[a]

[a]Universidade de São Paulo - USP, School of Dentistry, Department of Stomatology, São Paulo, SP, Brazil.

[b]Universidade de São Paulo - USP, School of Medicine, Department of Dermatology, São Paulo, SP, Brazil.

[c]Universidade de São Paulo - USP, School of Dentistry, Department of General Pathology, São Paulo, SP, Brazil.

[d]University of Western Ontario, Schulich School of Medicine & Dentistry, London, ON, Canada.

Declaration of Interests: The authors certify that they have no commercial or associative interest that represents a conflict of interest in connection with the manuscript.

Corresponding Author:
Camilla Vieira Esteves
E-mail: camilla.santos@usp.br

https://doi.org/10.1590/1807-3107bor-2019.vol33.0043

Submitted: September 17, 2018
Accepted for publication: November 11, 2018
Last revision: April 25, 2019

Abstract: Proteomic techniques have become popular in medicine and dentistry because of their widespread use in analyzing bodily fluids such as blood, saliva, urine, and gingival crevicular fluids as well as hard tissues such as enamel, dentine, and cementum. This review is a guide to proteomic techniques in general dentistry, summarizing techniques and their clinical application in understanding and diagnosing diseases and their use in identifying biomarkers of various diseases.

Keywords: Proteomics; Mass Spectrometry; Diagnosis; Electrophoresis, Polyacrylamide Gel; Saliva.

Introduction

Working with proteins, specifically with proteins from saliva, has been the focus of many research groups around the world during the last decade. Although the area is of widespread interest, many health professionals remain unaware of this technique or its applications. Here, therefore, we review the literature on the subject to update professionals in the field.

Proteomics—the study of the "proteome"—is widely used to analyze bodily fluids including blood, saliva, urine, sperm, gingival crevicular fluids, cervical-vaginal fluid and hard tissues such as enamel, dentine, and cementum, and is popular in both medicine and dentistry. With technological advances, and the utility of saliva as an indicator of systemic health that can be easily sampled using non-invasive methods,[1] with approximately 64% of human oral tissues studied to date.[2]

This review introduces use of proteomic analysis of saliva samples to general dentistry, summarizing its methodology, and clinical applications in studying disease.

What are proteins?

Proteins are diverse, genome-encoded biological macromolecules found in all cells. Proteins are covalently linked linear chains of combinations of 20 amino acids. Amino acids are the alphabet from which protein sequences are written; each amino acid has a side chain with distinct chemical properties,[3,4] meaning that different proteins made of different amino acid sequences have different properties. The number of possible combinations of amino acids explains how proteins can be the building blocks of all organisms

on earth, forming diverse components including enzymes, hormones, antibodies, transporters, spider webs, muscle fibers, and milk proteins[3]. This vast functional range, from a combination of simple building blocks, was captured by Sir Francis Crick: "*The most significant thing about proteins is that they can do almost anything*". "*Although proteins can act in so many ways, the way in which they are synthesized is probably uniform and rather simple, and this fits in with the modern view that gene action, being based on the nucleic acids, is also likely to be uniform and rather simple*".[5]

Siqueira and Dawes[6] believe that understanding protein function in the context of cells and bodily fluids is essential. Tools such as proteomic analysis are now contributing to our knowledge of this topic.

What is the proteome, and proteomics?

The concept of the "proteome" to describe a mixture of proteins was proposed by Wilkins et al.[7] and Wasinger et al.,[8] with the name being a *contraction of the "**PRO**teins" encoded by a given "gen**OME**", cell, or tissue type*.[7,8] A proteome has some differences from a genome[7] because it also takes into account post-transcriptional and post-translational process such as complex formation and covalent modification.[9]

Interestingly, the proteome changes under different conditions, such as physiological changes, or different stages of the cell cycle.[7,9] Proteomics profiling is the study of protein transcriptional profiles and interactions across all expressed proteins,[10] across cells, organisms, and even ecosystems, offering essential insights to understanding the function of proteins—and their roles in health.[1]

This review focuses on the use of mass spectrometry-based proteomics and can be divided into two fundamental strategies that provide different information on the proteome: bottom-up proteomics reports proteome composition, while top-down proteomics reports protein interactions and structure. In bottom-up proteomics, proteolytic digestion using trypsin yields protein fragments that can be identified using mass spectrometry, enabling inference of the proteins in the undigested sample (for example, shotgun gel-free proteomics).[11,12] Top-down proteomics uses mass spectrometry to analyze intact proteins, using direct protein fragmentation in gas-phase for insights into protein structure and complex formation; major examples are: High-Performance Liquid Chromatography-Mass Spectrometry (HPLC-MS), tandem mass-spectrometry (MS/MS) and matrix-assisted laser desorption/ionization time-of-flight MS/MS (MALDI-TOF MS/MS).[13]

Bottom-up proteomics is the most common approach. Bottom-up proteomics identifies proteins in a proteome by cross-referencing the masses of proteolytically-cleaved proteins measured by MS/MS to a pre-computed proteomic database. MS/MS is often preceded by separation using one-dimensional (1-D) or two-dimensional (2-D) gel electrophoresis, band excision, and in-gel digestion,[14] described in more detail below.

These fundamental proteomics strategies can be used for characterizing the proteome of a secretory gland, or whole saliva, and are capable of evaluating variation to distinguish normal and pathological proteomes toward identifying possible diseases biomarkers. Here we review sample purification, and describe ionization, mass analysis, and detection within the mass spectrometer for salivary proteomic analysis.

Protein purification techniques

To understand protein structure and function, it is first necessary to purify the protein. Various tools contribute to the purification process, including absorbance colorimetry, liquid chromatography, native and denaturing one-dimensional (1-D) and two-dimensional (2-D) gel electrophoresis, and western blotting.

Absorbance colorimetry measures light absorbance of a solute at specific wavelengths, allowing inference of the concentration of solute. Absorbance measurements enable the researcher to keep track of protein concentrations during the purification process

The core technique in protein purification is liquid chromatography, that separates proteins according to various properties, complementing later electrophoretic separation. Different types of liquid chromatography separate proteins as they pass through a column based on different protein properties. Anionic or cationic ion-exchange columns

specifically bind proteins based on charge; bound proteins are subsequently eluted by changing the ionic strength of the elution buffer. Size exclusion (also referred to as gel-filtration) columns separate proteins based on size, larger proteins are eluted first, while migration of smaller proteins is retarded by the column matrix. The matrix of affinity chromatography columns is decorated with chemical moieties; only specific proteins—usually due to genetic fusion of a moiety-binding affinity tag—bind these moieties, meaning that after washing other proteins from the column, an almost-pure sample of the tagged protein can be eluted from the column by adding a competing soluble chemical to the elution buffer. Liquid chromatography is usually performed using a High-Performance Liquid Chromatography (HPLC) pump device.[3] Individual liquid chromatography steps can be limited by co-elution of many different proteins, usually necessitating use of different consecutive techniques before acquisition of a pure protein sample.

Whether separated by chromatography, proteins are usually subsequently separated by native or denaturing gel electrophoresis in conjunction with western blotting to detect specific proteins using antibodies.

Protein separation and characterization by polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gel electrophoresis (PAGE) visualizes intact proteins, including isoforms of the same protein, at different positions on the gel.[13]

Either 1-D or 2-D PAGE can be used to separate components of the protein sample. One-dimensional PAGE separates proteins by molecular weight. Two-dimensional PAGE first separates proteins in 1-D by their isoelectric point (pI) using an immobilized pH gradient (IPG) strip that focuses each protein to the pH at which the protein has zero net charge; in the second dimension, the IPG strip is laid across the top of a polyacrylamide gel for electrophoretic separation by size, as in 1-D gel electrophoresis.[15,16,17]

The protocol chosen for 2-D PAGE should optimally solubilize all proteins analyzed, remove contamination, prevent protein aggregation during focusing, and avoid chemical modification of the sample. Proteins are denatured and solubilized using high concentrations of urea; thiourea can be added to further increase the solution's solubilization ability. Reducing agents are typically added to reduce disulfide bonds, although the most commonly used reducing agent, dithiothreitol (DTT), is responsible for "point streaking" artefacts in 2-D gels, depending on the sample,[18] which must be considered during protocol development.

Due to its ability to resolve approximately 5,000 different proteins, depending on the size of gel,[18,19] 2-D PAGE continues to be widely used to study the proteome, and can successfully characterize post-translational modifications and mutant proteins.

Protein separation according to molecular weight is based on pore sizes within the polyacrylamide gel. Those pores can be modified to optimally separate proteins of interest by changing the acrylamide percentage when making the gel.

Sodium dodecyl sulfate (SDS)-PAGE is a type of 1-D PAGE that enables separation of proteins based purely on their molecular weight. SDS-PAGE involves addition of SDS when making the gel, and heat-denaturing the protein sample (i.e., unfolding their native structural conformation) by disrupting non-covalent bonds. Negatively-charged SDS molecules coat the unfolded protein in proportion to the protein's molecular weight; thus protein SDS-PAGE separation is not influenced by protein structure, and is directly proportional to protein molecular weight alone.

Proteins are visualized in the gel after electrophoresis by staining. Coomassie blue is the best-known protein staining technique. Because Coomassie blue's detection limit is around 100 ng, however, small proteins, even when abundant, are difficult to visualize. In cases involving low proteins concentrations, high-sensitivity silver staining, capable of detecting less than 1 ng protein, can be used.[18] Silver staining methods using aldehyde-based fixatives/sensitizers, however, prevent subsequent protein analysis by mass spectrometry (MS) due to protein cross-linkage.[19]

SDS-PAGE is an analytical technique incapable of direct identification of the separated proteins. Proteins(s) of interest can be identified by specific antibodies in a western blot, or by gel extraction, proteolysis, and MS or MS/MS analysis.

Protein separation and characterization by native gels

Although SDS-PAGE is the most common form of PAGE protein separation, native PAGE of the natively-folded protein can also be performed without SDS.[20] Because native PAGE separates proteins according to both their charge and hydrodynamic size, it provides complementary information to an SDS-PAGE gel. Because native PAGE is a kind of 1-D electrophoresis that separates proteins by charge, protein pI and electrophoresis pH will influence protein mobility.

Protein identification by western Blot

Western blotting using a specific antibody can locate a specific protein in a gel. Proteins are first separated by PAGE, the gel transferred onto a membrane, incubated with a specific antibody, and antibody location visualized.[3,21] In the western blot analysis, the passage of gel electrophoresis is included to solve the cross-reactivity problem of the antibodies.[22]

The amount of protein in a sample that will be study by this technique can be quantified for comparison between groups.

Methodology

Protein Ionization techniques

For proteomic analysis, as many proteins as possible in the sample must be identified using mass spectrometry. Mass spectrometry requires that proteins in solution or solid state are ionized in gas phase for injection and acceleration in an electric or magnetic field for analysis. The two main ionization methods that minimize sample fragmentation are MALDI (Figure 1A) and ESI[23] (Figure 1B). ESI is used for aqueous solutions, in which a syringe containing the sample is forced through a hypodermic needle connected to a high-voltage power supply. The high voltage induces an intense electric field at the needle tip that disperses the sample into a fine spray of charged droplets; the solvent subsequently evaporates in a warmed chamber, leaving desolvated sample ions.[24,25] MALDI, on the other hand, involves embedding the sample in an organic matrix that is sublimed by pulses of a ultraviolet (UV) laser, forming gaseous peptide or protein ions for MS analysis.[25,26]

Mass analysis of proteins

After ionization by MALDI or ESI, ion mass-to-charge (m/z) ratios are measured by the mass spectrometer. Protein mass analysis is performed either "in space" or "in time", enabling recording over a wide range of masses.[20,23] In space m/z measurements are performed sequentially through the travel of ions into an instrument such as triple quadruple (TQ), quadrupole/time-of-flight (Q-TOF),

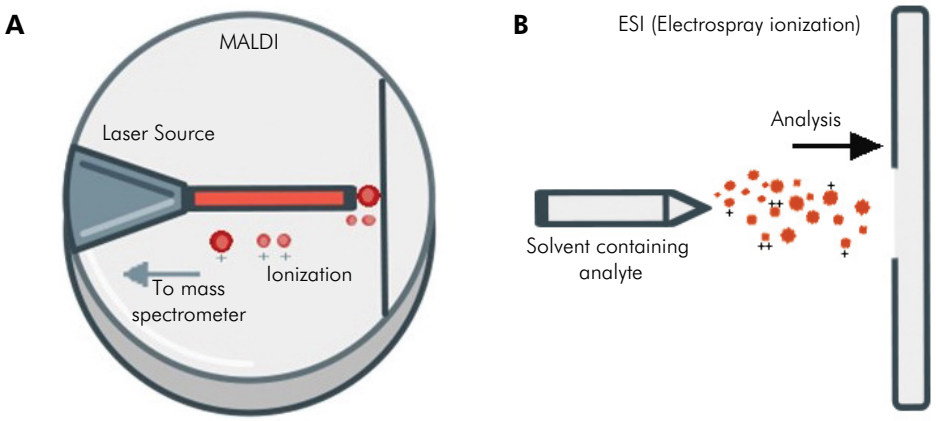

Figure 1. A: Ionization methods - MALDI: the peptides are crystallizing into an organic matrix and are ionization by laser; B: Ionization methods - ESI: The aqueous solution containing the analyte is forced through a capillary needle which upon evaporation of the solute generates ionized form.

or time-of-flight/time-of-flight (TOF-TOF). In time analyses use the same analyzer, such as quadrupole ion trap (Q-IT).[26]

The most widely used instruments for peptide mass analysis are time-of-flight (TOF) analyzers. The ion is accelerated with speed inversely proportional to mass by the potential difference between electrodes in a vacuum tube; the time of detection is therefore proportional to m/z. The time of flight can be used to calculate the m/z value, and a plot of m/z versus intensity (ion count), or MS spectrum, generated for interpretation on a workstation.[27,28] Each m/z peak in the MS spectrum is subsequently annotated by cross-referencing to a database such as MASCOT[29] or SEQUEST.

Tandem Mass Spectrometry (MS/MS) for protein identification

Tandem Mass Spectrometry, popularly known such as MS/MS, is a technique that can deduce protein sequences based on patterns of protein fragmentation in a collision chamber located between two mass spectrometers[30]. Sample ions are introduced into the first mass spectrometer as described above, separated by charge, and selected m/z ions directed to the collision chamber for fragmentation by collision with neutral atoms or molecules. Resulting peptide fragments are directed to the second mass spectrometer for m/z analysis. The tandem mass spectra represent the peptides derived from successive fragmentations, and, can be used to deduce the precise sequence of amino acids in the sample[26] (Figure 2).

Data analysis to identify proteins

Integral to MS spectrum annotation is cross-referencing observed m/z peaks with established databases. The most popular databases, MASCOT and SEQUEST,[31,32] are used to determine which peptide sequence in a protein database corresponds to each m/z peak.[32] SEQUEST, the first major database-based search algorithm established in 1994, uses cross-correlation to analyze the similarity between experimental and theoretical mass spectra. The other major software, MASCOT, introduced in 1999, is based on a probability-based score.[26,33] MASCOT is universally used by research facilities around the world and uses a probabilistic scoring algorithm for protein identification that was adapted from the MOWSE algorithm.[34] MASCOT is freely available but acquiring a license introduces additional features.[29,35] For the best match of the protein sequence, MASCOT employs the widely used significance level of 0.05.

The SEQUEST database is typically used when trypsin is used for protein digestion, and peptides are separated by liquid chromatography before proceeding to MS/MS. This procedure producing several MS/MS spectra and the SEQUEST database was the first software to fulfill this need.[9,36]

Salivary proteome databases

Approximately 60% (11,716) of all human proteins (19,613) are expressed in the salivary gland, and of these, the expression of 85 is elevated compared with other tissue types[37] as demonstrated by recent development of the Human Protein Atlas

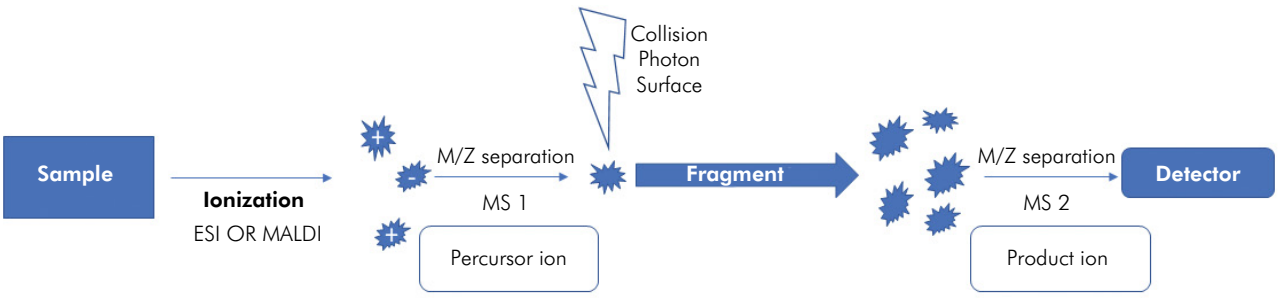

Figure 2. Tandem Mass Spectrometry steps

portal (www.proteinatlas.org), an interactive, frequently-updated database that catalogues proteins predicted to be expressed based on RNA level.[38]

Work on salivary proteomics and bioinformatic analysis tools has developed significantly in recent years, enabling cataloging of proteins found in saliva, relationship of the salivary proteome to other proteomes, and highlighting different proteome profiles for different salivary glands. A 2009 study that focused on fractionating the salivary proteome identified 2340 salivary proteins, of which 20% matched plasma proteome proteins;[39] a subsequent review by Amado et al.[40] in 2013 listed more than 3,000 different salivary proteins. Interestingly, this database identified proteins expressed at elevated levels in salivary glands and created a map of differential salivary gland protein expression from ductal epithelial cells, and serous or mucinous cells. According to the Human Protein Atlas, and confirmed using antibody based profiling and proteogenomics,[38] four examples of proteins more expressed in serous salivary glands are CA6, CST2 and PIP and AMY1B. One example of a protein specific for mucinous salivary glands is MUC7, and one example expressed in salivary ducts is SLC5A5.

Another important information to better understand the finds in mass-spectrometry is the database for updated information about protein sequence, functions, and annotation data is The Universal Protein Resource - UniProt (https://www.uniprot.org). The UniProt databases are a combination of the UniProt Knowledgebase (UniProtKB), UniProt Reference Clusters (UniRef), and the UniProt Archive (UniParc). UniProt emerged as a collaborative online platform between the European Bioinformatics Institute (EMBL-EBI), the SIB (Swiss Institute of Bioinformatics) and the Protein Information Resource (PIR), that are together responsible for database curation, software development and support.

Secreted salivary proteins

Around 400 salivary proteins have been identified by gel electrophoresis[1,40–42]; the most abundant are mucins, proline rich proteins (PRPs), statherin, histatins, amylase, and lysozyme:

a. Five types of mucins identified in the oral cavity are MUC5B, MUC7, MUC19, MUC1, and MUC4, each composed of a unique domain structure that influences its properties.[43] Salivary mucins are produced in submandibular, sublingual, and minor salivary mucous glands to bind and retain antibacterial proteins in the oral cavity to prevent T-cell and epithelial cell infections.[13,43,44,45]
b. PRPs, encoded by PRB1, are expressed in the serous gland, and are subdivided into acidic (aPRPs), basic (bPRPs) and basic-glycosylated (gPRPs).[11,46] The main function of aPRPs is bind and maintain the concentration of calcium in saliva.
c. Statherin maintains a supersaturated level of calcium in the saliva by preventing its precipitation. Statherin helps remineralization of teeth.[13,47] Its concentration is not influenced to circadian rhythms, that could be very interesting, one time that the period of collection did not interfere in the result.
d. Histatins are antibacterial and antifungal proteins that are among the most abundant salivary proteins; histatins 1,3 and 5 were first described by Oppenheim et al.[48] in 1988 who linked their presence to anti-fungal activity against *Candida albicans*. Consistent with this first study, Siqueira et al.[49] found histatins within the acquired enamel pellicle that may protect against acid damage; *in vitro* studies by Moffa et al.[50] suggest that a histatin 5 coat on epithelial cell surfaces reduces *C. albicans* colonization; and Siqueira et al.[51] used a proteomic approach to identify 43 complexes of histatin 1 with other salivary proteins, finding that the killing activity of a histatin 1/amylase complex is better than histatin 1 alone, for example.
e. Salivary amylase, the most abundant secreted protein in parotid saliva, catalyzes hydrolysis of starch into sugars.[52,53] There are five human amylase isoenzymes and three are found in saliva can be assigned to family A and family B, based on differences in post-translational glycosylation content.[53] Glycosylated Family

A has molecular weight of 62 kDa whereas nonglycosylated Family B has a molecular weight of 56kDa.[54] Due to its abundance in saliva, amylase may plan an additional role in maintaining oral mucosa. A recent proteomics study analyzing the interactome of amylase in whole saliva[55] confirmed that amylase interacts with mucins, although histatins were not detected likely due to their short lifespan in whole saliva. These findings suggested that amylase both protects and functionally modulates its partners in addition to digesting carbohydrates.

Clinical applications of salivary proteomics

Proteomic approaches were initially used to characterize all proteins in a given cell. As the technique developed, however, the goal shifted to detect differences in proteomes related to disease. The proteomic approach allowed control the gene transcriptions through the interconnection of the extracellular microenvironment that characterizes the information flow over protein pathways. Future progress will enable understanding the course of the disease and identification of disease biomarkers, enabling early detection, and targeted, patient-tailored therapy.

Clinical research using proteomics is growing. A review by Colantonio et al.[56] on clinical application of proteomics in 2004 found 192 articles; In an update at 2018 made for this review, our group found 14,425 articles, been with the rose of publication in this area. This update confirmed the impressions of Colantonio et al.[56] in 2004, who wrote "*These numbers suggest that the application of clinical proteomic research is growing rapidly in the field of biomarker discovery*".

Saliva as a potential source of disease biomarkers

Many diseases could be better understood by advances in proteomics. Saliva is an excellent medium in which to detect disease biomarkers because it is easy and painless to collect, and contains proteins rich in information about diseases process.

Saliva is an important fluid that lubricates oral mucosa and ingested food, protects and maintains tissue integrity, neutralizes acids from bacteria and food, and cleans the oral cavity (Figure 3).[14,47,57] Whole saliva derives from major salivary glands comprising two parotid, two submandibular, and two sublingual glands; minor salivary glands; and crevicular fluid, in total producing between 600 and 12,000 mL/day.[58]. Saliva is an exocrine secretion that can be divided into mucous and serous components. The parotid gland secretes only serous component, the submandibular gland secretes a mixture of mucous and serous components, while the sublingual and minor glands secrete only mucous components.[57,59]

Use of saliva is attractive for monitoring health and disease because its collection, is non-invasive, easy, and painless, and does not require special training. Recent advances in proteomics enables it to qualitatively and quantitatively determine the protein composition of saliva.[58] Indeed, In the last decade, advances in saliva research have identified many proteins as potential systemic biomarkers for endocrine function, stress and psychological state, exposure to infectious agents, use or metabolism of drugs or other xenobiotics, and cancers.[59]

Siqueira and Dawes[6] illustrated how saliva is better than plasma for detection of biomarkers: while the 22 most abundant proteins in plasma make up 99% of the total plasma protein content, making identification of the remainder of the 2676 serum proteins a challenge, the 20 most abundant salivary proteins make up only 40% of the total salivary protein content, facilitating ease of detection of biomarkers from the remaining 60% of proteins.

More studied are needed to compare healthy and diseased salivary proteomes. While Khurshid et al.[1] demonstrated that approximately "*30% of proteomic studies of saliva explored the proteome during oral pathological conditions such as caries, periodontitis, gingivitis, dental abscess, endodontic lesions, and oral carcinomas*", there are few studies of the salivary proteome under normal conditions that would provide a gold standard reference for other studies, principally of oral disease conditions.

Most studies have used whole saliva instead of specific glandular secretions to discover possible oral biomarkers. Because whole saliva is composed of glandular secretions (95.6%) mixed with gingival

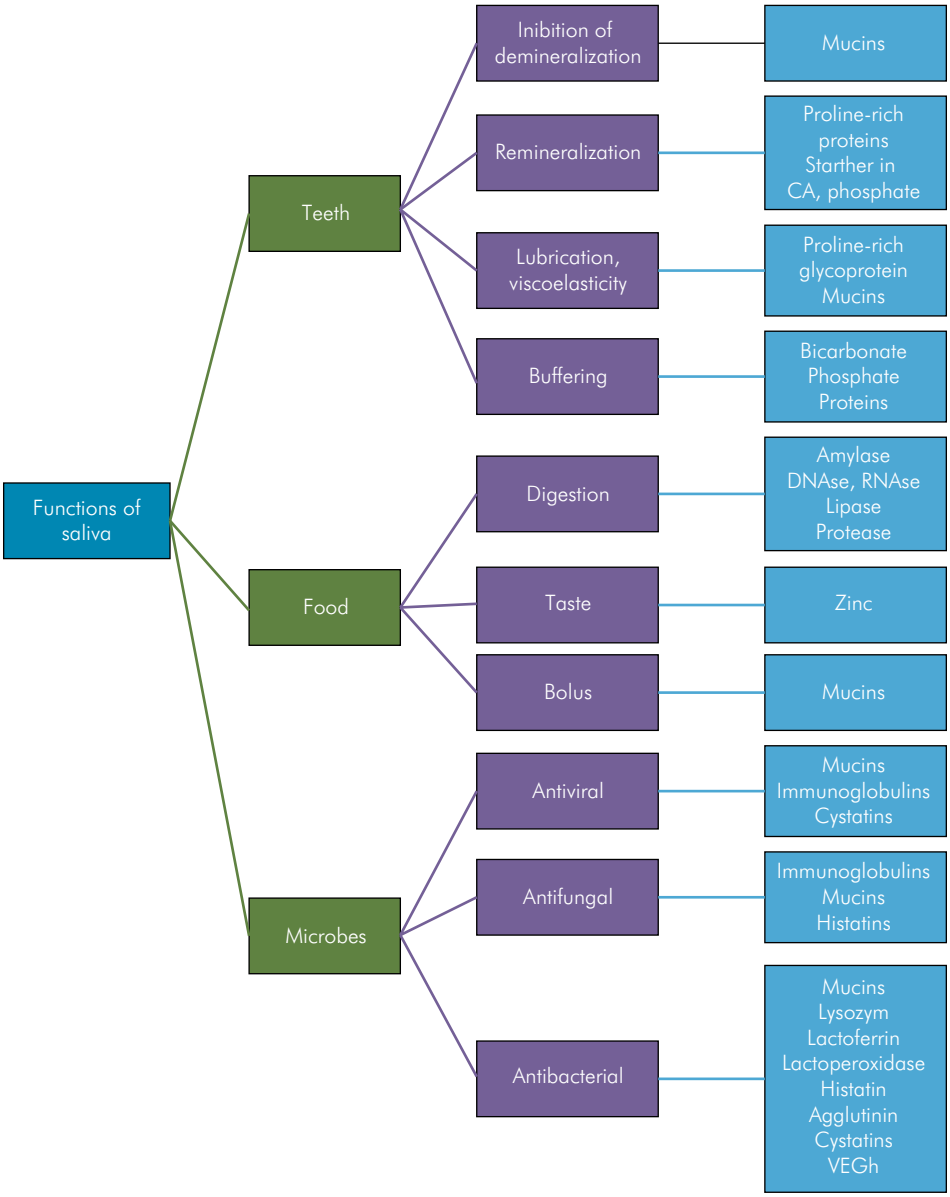

Figure 3. Saliva functions in the oral cavity

fluid (2.4%), microorganisms (1%), and epithelial host cells (1%), samples may contain additional proteins from, for example, gingival inflammation discharge. Isolated glandular secretions obtained through cannulation of major gland ducts may prove better for biomarker detection, avoiding the influence of the other whole saliva components.[6]

Nevertheless, many studies reports use of oral fluids such as whole saliva for early detection and monitoring of diseases. We used the terms "saliva" and "proteomic" to search PubMed and Google Scholar for such studies, summarizing 16 highly-cited articles according to citations in Google Scholar in Table 1. The number of publications per year for "saliva" and "proteomic" (based on https://www.ncbi.nlm.nih.gov/pubmed/) is shown in Figure 4.

Many authors believe that comprehensive characterization of human saliva is an important step toward the clinical application of salivary proteomics for the diagnosis and prognosis of diseases.

Table 1. Highly-cited articles on disease detection using salivary proteomics.

Year	Author	Salivary proteomics application	Results
2007	Hu et al.[14]	Detection of oral cancer and Sjögren syndrome	Slight advantage over serum in detection
2007	Giusti et al.[60]	Detection of Sjögren syndrome	Proteomics may be useful in diagnosis of connective tissue disorders
2008	Hu et al.[61]	Detection of oral cancer	Promising approach to searching for oral cancer biomarkers
2009	Wu et al.[62]	Detection of gastric cancer	Promising approach for early clinical diagnosis of gastric cancer
2009	Wu et al.[63]	Profile of generalized aggressive periodontitis	Proteome analysis may contribute to understanding the etiology of aggressive periodontitis
2009	Yan et al.[64]	Systematic comparison of the human saliva and plasma proteomes	High similarity between saliva and plasma proteomes
2010	de Jong et al.[65]	Detection of oral cancer	Promising detection method for oral cancer
2010	Wu et al.[66]	Potential biomarkers in saliva for oral squamous cell carcinoma	Salivary screening can be the best option for primary screening of high-risk cases of oral cancer
2010	Gonçalves et al.[67]	Analysis of chronic periodontitis patients	Proteome analysis may contribute to the improvement of periodontal diagnosis
2011	Castagnola et al.[59]	Potential applications of human saliva as a diagnostic fluid	A proteomic analysis approach may be useful in diagnosis of head and neck cancers, breast and gastric cancers, salivary gland function and disease, Sjögren syndrome, systemic sclerosis, dental and gingival pathology, systemic, psychiatric, and neurological diseases
2011	Xiao et al.[68]	Biomarkers discovery in human saliva	Found salivary proteomic biomarkers for oral cancer, Sjögren syndrome and breast cancer
2012	Xiao et al.[69]	Detection of lung cancer	Proteomic biomarkers are present in human saliva when people develop lung cancer
2012	Ambatipudi et al.[70]	Detection of primary Sjögren syndrome	Proteome analysis improved and provided early diagnosis of primary Sjögren syndrome
2013	Amado et al.[40]	Overview of the major achievements in saliva proteomics	Identified protein markers for chronic periodontitis, gingivitis, head and neck squamous cell carcinoma, non-invasive breast cancer, rheumatoid arthritis, periodontitis, aggressive periodontitis, dental caries, Sjögren syndrome, lung cancer, type 2 diabetes in the elderly, orthodontic treatment, pre-malignant and malignant lesions, gastric cancer, periodontitis with obesity, oral leukoplakia, type 1 diabetes,
2013	Martins et al.[71]	Systematic review of salivary proteins as biomarkers for dental caries	Lack of sufficient evidence to establish salivary proteins as a biomarker for this disease
2017	Zuanazzi et al.[72]	Postnatal identification of Zika virus peptides from saliva	Possibility of non-invasive sampling of saliva for subsequent large-scale screening for Zika virus infection

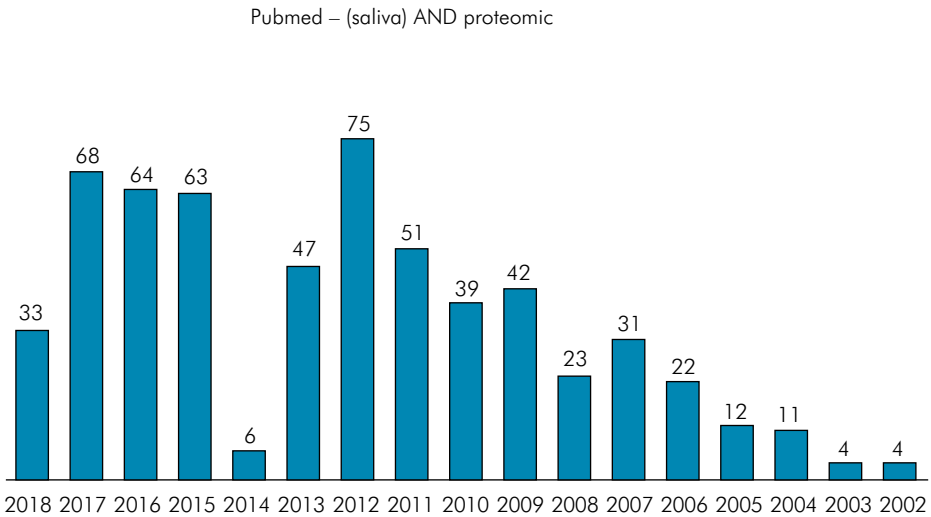

Figure 4. Publications numbers increase according to advantages in proteomics technique, increasing the possibility of discovering new diseases biomarkers. The rose in number was at 2012 with 75 new published studies in saliva area.

Conclusion

In this review we have highlighted how salivary proteomic analysis has broad practical promise and application for easy and painless diagnosis of diseases. The review also outlined to the general dentist a step-by-step summary of how salivary proteomics can be used for diagnosis, and search for biomarkers. Despite salivary proteomics being a relatively new area of research, recent advances, and the increase in published studies, demonstrate the great potential of saliva as a diagnostic fluid.

Acknowledgements

The authors acknowledge the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/ CAPES for the research grants (PDSE – Edital nº19/2016.).

References

1. Khurshid Z, Zohaib S, Najeeb S, Zafar MS, Rehman R, Rehman IU. Advances of proteomic sciences in dentistry. Int J Mol Sci. 2016 May;17(5):728. https://doi.org/10.3390/ijms17050728
2. Maron JL. Salivary diagnostics in pediatrics : applicability, translatability, and limitations. Front Public Health..2017 Apr;5:83. https://doi.org/10.3389/fpubh.2017.00083
3. Nelson DL, Cox MM. Lehninger principles of biochemistry. 5th ed. New York: WH Freeman; 2008.
4. Breda A, Valadares NF, Souza ON, Garratt RC. Protein structure, modelling and applications. In: Gruber A, Durham AM, Huynh C, Portillo HA, eds. Bioformatics in tropical disease research: a pratical and case-study approach. Bethesda: National Center for Biotechnology Information; 2008.
5. Crick FH. On protein synthesis. Symp Soc Exp Biol. 1958;12:138-63.
6. Siqueira WL, Dawes C. The salivary proteome: challenges and perspectives. Proteomics Clin Appl. 2011 Dec;5(11-12):575-9. https://doi.org/10.1002/prca.201100046
7. Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, et al. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol Genet Eng Rev. 1996;13(1):19-50. https://doi.org/10.1080/02648725.1996.10647923
8. Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, et al. Progress with gene-product mapping of the Mollicutes: mycoplasma genitalium. Electrophoresis. 1995 Jul;16(7):1090-4. https://doi.org/10.1002/elps.11501601185
9. Barbosa EB, Vidotto A, Polachini GM, Henrique T, Marqui AB, Tajara EH. Proteomics: methodologies and applications to the study of human diseases. Rev Assoc Med Bras. 2012;58(3):366-75. https://doi.org/10.1590/S0104-42302012000300019
10. Mrowka R, Patzak A, Herzel H. Is there a bias in proteome research? Genome Res. 2001 Dec;11(12):1971-3. https://doi.org/10.1101/gr.206701
11. Messana I, Cabras T, Inzitari R, Lupi A, Zuppi C, Olmi C, Fadda MB, Cordaro M, et al. Characterization of the human salivary basic proline-rich protein complex by a proteomic approach. J Proteome Res. 2004 Jul-Aug;3(4):792-800. https://doi.org/10.1021/pr049953c
12. Bogdanov B, Smith RD. Proteomics by FTICR mass spectrometry: top down and bottom up. Mass Spectrom Rev. 2005 Mar-Apr;24(2):168-200. https://doi.org/10.1002/mas.20015
13. Cabras T, Iavarone F, Manconi B, Olianas A, Sanna MT, Castagnola M, et al. Top-down analytical platforms for the characterization of the human salivary proteome. Bioanalysis. 2014 Feb;6(4):563-81. https://doi.org/10.4155/bio.13.349
14. Hu S, Loo JA, Wong DT. Human saliva proteome analysis. Ann N Y Acad Sci. 2007 Mar;1098(1):323-9. https://doi.org/10.1196/annals.1384.015
15. Aslam B, Basit M, Nisar MA, Khurshid M, Rasool MH. Proteomics: Technologies and Their Applications. J Chromatogr Sci. 2017 Feb;55(2):182-96. https://doi.org/10.1093/chromsci/bmw167
16. Washburn MP, Wolters D, Yates 3rd JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol. 2001 Mar;19(3):242-7. https://doi.org/10.1038/85686
17. Dowsey AW, Dunn MJ, Yang GZ. The role of bioinformatics in two-dimensional gel electrophoresis. Proteomics.. 2003 Aug;3(8):1567-96. https://doi.org/10.1002/pmic.200300459
18. Weiss W, Weiland F. Görg A. Protein detection and quantitation technologies for gel-based proteome analysis. Methods Mol Biol. 2009;564:59-82. https://doi.org/10.1007/978-1-60761-157-8_4

19. Görg A, Weiss W, Dunn MJ. Current two-dimensional electrophoresis technology for proteomics. Proteomics. 2004 Dec;4(12):3665-85. https://doi.org/10.1002/pmic.200401031
20. May C, Brosseron F, Chartowski P, Schumbrutzki C, Schoenebeck B, Marcus K. Instruments and methods in proteomics. Methods Mol Biol. 2011;696:3-26. https://doi.org/10.1007/978-1-60761-987-1_1
21. Mahmood T, Yang PC. Western blot: technique, theory, and trouble shooting. N Am J Med Sci. 2012 Sep;4(9):429-34. https://doi.org/10.4103/1947-2714.100998
22. Towbin H, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA. 1979 Sep;76(9):4350-4. https://doi.org/10.1073/pnas.76.9.4350
23. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003 Mar;422(6928):198-207. https://doi.org/10.1038/nature01511
24. Nguyen S, Fenn JB. Gas-phase ions of solute species from charged droplets of solutions. Proc Natl Acad Sci USA. 2007 Jan;104(4):1111-7. https://doi.org/10.1073/pnas.0609969104
25. Chait BT. Mass spectrometry in the postgenomic era. Annu Rev Biochem. 2011;80(1):239-46. https://doi.org/10.1146/annurev-biochem-110810-095744
26. Hernandez P, Müller M, Appel RD. Automated protein identification by tandem mass spectrometry: issues and strategies. Mass Spectrom Rev. 2006 Mar-Apr;25(2):235-54. https://doi.org/10.1002/mas.20068
27. Hrabák J, Chudácková E, Walková R. Matrix-assisted laser desorption ionization-time of flight (maldi-tof) mass spectrometry for detection of antibiotic resistance mechanisms: from research to routine diagnosis. Clin Microbiol Rev. 2013 Jan;26(1):103-14. https://doi.org/10.1128/CMR.00058-12
28. Barbosa EB, Vidotto A, Polachini GM, Henrique T, Marqui ABT, Tajara EH. Proteomics: methodologies and applications to the study of human diseases. Rev Assoc Med Bras. 2012 May;58(3):366-75. https://doi.org/10.1590/S0104-42302012000300019
29. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999 Dec;20(18):3551-67. https://doi.org/10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2
30. Niessen WM, Sn. In M. Encyclopedia of Spectroscopy and Spectrometry. Elsevier; 1999. pp. 1675-81. https://doi.org/10.1016/B978-0-12-374413-5.00215-3
31. Cottrell JS. Protein identification using MS/MS data. J Proteomics. 2011 Sep;74(10):1842-51. https://doi.org/10.1016/j.jprot.2011.05.014
32. Zhang W, Zhao X. Method for rapid protein identification in a large database. BioMed Res Int. 2013;2013:414069. https://doi.org/10.1155/2013/414069
33. Wenger CD, Coon JJ. A proteomics search algorithm specifically designed for high-resolution tandem mass spectra. J Proteome Res. 2013 Mar;12(3):1377-86. https://doi.org/10.1021/pr301024c
34. Pappin DJ, Hojrup P, Bleasby AJ. Rapid identification of proteins by peptide-mass fingerprinting. Curr Biol. 1993 Jun;3(6):327-32. https://doi.org/10.1016/0960-9822(93)90195-T
35. Koenig T, Menze BH, Kirchner M, Monigatti F, Parker KC, Patterson T, et al. Robust prediction of the MASCOT score for an improved quality assessment in mass spectrometric proteomics. J Proteome Res. 2008 Sep;7(9):3708-17. https://doi.org/10.1021/pr700859x
36. Eng JK, Mccormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom. 1994 Nov;5(11):976-89. https://doi.org/10.1016/1044-0305(94)80016-2
37. The Human Protein Atlas. Downloadable data. 2018 [cited 2018 Mar 23]. Available from: https://www.proteinatlas.org/about/download
38. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419. https://doi.org/10.1126/science.1260419
39. Bandhakavi S, Stone MD, Onsongo G, Van Riper SK, Griffin TJ. A dynamic range compression and three-dimensional peptide fractionation analysis platform expands proteome coverage and the diagnostic potential of whole saliva. J Proteome Res. 2009 Dec;8(12):5590-600. https://doi.org/10.1021/pr900675w
40. Amado FM, Ferreira RP, Vitorino R. One decade of salivary proteomics: current approaches and outstanding challenges. Clin Biochem. 2013 Apr;46(6):506-17. https://doi.org/10.1016/j.clinbiochem.2012.10.024
41. Hu S, Xie Y, Ramachandran P, Ogorzalek Loo RR, Li Y, Loo JA, et al. Large-scale identification of proteins in human salivary proteome by liquid chromatography/mass spectrometry and two-dimensional gel electrophoresis-mass spectrometry. Proteomics. 2005 Apr;5(6):1714-28. https://doi.org/10.1002/pmic.200401037
42. Vitorino R, Lobo MJ, Ferrer-Correira AJ, Dubin JR, Tomer KB, Domingues PM, et al. Identification of human whole saliva protein components using proteomics. Proteomics. 2004 Apr;4(4):1109-15. https://doi.org/10.1002/pmic.200300638
43. Frenkel ES, Ribbeck K. Salivary mucins in host defense and disease prevention. J Oral Microbiol. 2015 Dec;7(1):29759. https://doi.org/10.3402/jom.v7.29759
44. Ma J, Rubin BK, Voynow JA. Mucins, mucus, and goblet cells. Chest. 2018 Jul;154(1):169-176. https://doi.org/10.1016/j.chest.2017.11.008

45. Chen Y, Zhao YH, Kalaslavadi TB, Hamati E, Nehrke K, Le AD, et al. Genome-wide search and identification of a novel gel-forming mucin MUC19/Muc19 in glandular tissues. Am J Respir Cell Mol Biol. 2004 Feb;30(2):155-65. https://doi.org/10.1165/rcmb.2003-0103OC
46. Bennick A. Salivary proline-rich proteins. Mol Cell Biochem. 1982 Jun;45(2):83-99. https://doi.org/10.1007/BF00223503
47. Khurshid Z, Zohaib S, Najeeb S, Zafar MS, Slowey PD, Almas K. Human Saliva Collection Devices for Proteomics: an Update. Int J Mol Sci. 2016 Jun;17(6):846. https://doi.org/10.3390/ijms17060846
48. Oppenheim FG, Xu T, McMillian FM, Levitz SM, Diamond RD, Offner GD, et al. Histatins, a novel family of histidine-rich proteins in human parotid secretion. J Biol Chem. 1988 Jun 5;263(16):7472-7.
49. Siqueira WL, Margolis HC, Helmerhorst EJ, Mendes FM, Oppenheim FG. Evidence of intact histatins in the in vivo acquired enamel pellicle. J Dent Res. 2010 Jun;89(6):626-30. https://doi.org/10.1177/0022034510363384
50. Moffa EB, Machado MA, Mussi MC, Xiao Y, Garrido SS, Giampaolo ET, et al. In vitro identification of histatin 5 salivary complexes. PLoS One. 2015 Nov 6;10(11):e0142517. https://doi.org/10.1371/journal.pone.0142517
51. Siqueira WL, Lee YH, Xiao Y, Held K, Wong W. Identification and characterization of histatin 1 salivary complexes by using mass spectrometry. Proteomics. 2012 Nov;12(22):3426-35. https://doi.org/10.1002/pmic.201100665
52. Loo JA, Yan W, Ramachandran P, Wong DT. Comparative human salivary and plasma proteomes. J Dent Res. 2010;89(10):1016-23. https://doi.org/10.1177/0022034510380414
53. Boehlke C, Zierau O, Hannig C. Salivary amylase - The enzyme of unspecialized euryphagous animals. Arch Oral Biol. 2015 Aug;60(8):1162-76. https://doi.org/10.1016/j.archoralbio.2015.05.008
54. Keller PJ, Kauffman DL, Allan BJ, Williams BL. Structural differences between the isoenzymes of human parotid α-amylase. Biochemistry. 1971 Dec;10(26):4867-74. https://doi.org/10.1021/bi00802a006
55. Crosara KT, Zuanazzi D, Moffa EB, Xiao Y, Machado MA, Siqueira WL. Revealing the amylase interactome in whole saliva using proteomic approaches. BioMed Res Int. 2018 Jan;2018:6346954. https://doi.org/10.1155/2018/6346954
56. Colantonio DA, Chan DW. The clinical application of proteomics. Clin Chim Acta. 2005 Jul;357(2):151-8. https://doi.org/10.1016/j.cccn.2005.03.020
57. Proctor GB, Carpenter GH. Salivary secretion: mechanism and neural regulation. Monogr Oral Sci. 2014;24:14-29. https://doi.org/10.1159/000358781
58. Wong D. Salivary diagnostics. Ames: Wiley-Blackwell; 2008.
59. Castagnola M, Picciotti PM, Messana I, Fanali C, Fiorita A, Cabras T, et al. Potential applications of human saliva as diagnostic fluid. Acta Otorhinolaryngol Ital. 2011 Dec;31(6):347-57.
60. Giusti L, Baldini C, Bazzichi L, Ciregia F, Tonazzini I, Mascia G, et al. Proteome analysis of whole saliva: a new tool for rheumatic diseases—the example of Sjögren's syndrome. Proteomics. 2007 May;7(10):1634-43. https://doi.org/10.1002/pmic.200600783
61. Hu S, Arellano M, Boontheung P, Wang J, Zhou H, Jiang J, et al. Salivary proteomics for oral cancer biomarker discovery. Clin Cancer Res. 2008 Oct;14(19):6246-52. https://doi.org/10.1158/1078-0432.CCR-07-5037
62. Wu ZZ, Wang JG, Zhang XL. Diagnostic model of saliva protein finger print analysis of patients with gastric cancer. World J Gastroenterol. 2009 Feb;15(7):865-70. https://doi.org/10.3748/wjg.15.865
63. Wu Y, Shu R, Luo LJ, Ge LH, Xie YF. Initial comparison of proteomic profiles of whole unstimulated saliva obtained from generalized aggressive periodontitis patients and healthy control subjects. J Periodontal Res. 2009 Oct;44(5):636-44. https://doi.org/10.1111/j.1600-0765.2008.01172.x
64. Yan W, Apweiler R, Balgley BM, Boontheung P, Bundy JL, Cargile BJ, et al. Systematic comparison of the human saliva and plasma proteomes. Proteomics Clin Appl. 2009 Jan;3(1):116-34. https://doi.org/10.1002/prca.200800140
65. de Jong EP, Xie H, Onsongo G, Stone MD, Chen XB, Kooren JA, et al. Quantitative proteomics reveals myosin and actin as promising saliva biomarkers for distinguishing pre-malignant and malignant oral lesions. PLoS One. 2010 Jun;5(6):e11148. https://doi.org/10.1371/journal.pone.0011148
66. Wu JY, Yi C, Chung HR, Wang DJ, Chang WC, Lee SY, et al. Potential biomarkers in saliva for oral squamous cell carcinoma. Oral Oncol. 2010 Apr;46(4):226-31. https://doi.org/10.1016/j.oraloncology.2010.01.007
67. Gonçalves LDR, Soares MR, Nogueira FCS, Garcia C, Camisasca DR, Domont G, et al. Comparative proteomic analysis of whole saliva from chronic periodontitis patients. J Proteomics. 2010 May;73(7):1334-41. https://doi.org/10.1016/j.jprot.2010.02.018
68. Xiao H, Wong DT. Proteomics and its applications for biomarker discovery in human saliva. Bioinformation. 2011 Jan;5(7):294-6. https://doi.org/10.6026/97320630005294
69. Xiao H, Zhang L, Zhou H, Lee JM, Garon EB, Wong DTW. Proteomic analysis of human saliva from lung cancer patients using two-dimensional difference gel electrophoresis and mass spectrometry. Mol Cell Proteomics. 2012;11(2):M111.012112. https://doi.org/10.1074/mcp.M111.012112
70. Ambatipudi KS, Swatkoski S, Moresco JJ, Tu PG, Coca A, Anolik JH, et al. Quantitative proteomics of parotid saliva in primary Sjögren's syndrome. Proteomics. 2012 Oct;12(19-20):3113-20. https://doi.org/10.1002/pmic.201200208

71. Martins C, Buczynski AK, Maia LC, Siqueira WL, Castro GF. Salivary proteins as a biomarker for dental caries: a systematic review. J Dent. 2013 Jan;41(1):2-8. https://doi.org/10.1016/j.jdent.2012.10.015
72. Zuanazzi D, Arts EJ, Jorge PK, Mulyar Y, Gibson R, Xiao Y, et al. Postnatal Identification of Zika Virus Peptides from Saliva. J Dent Res. 2017 Sep;96(10):1078-84. https://doi.org/10.1177/0022034517723325