Acta Scientific Dental Sciences (ISSN: 2581-4893)

Volume 5 Issue 4 April 2021

Research Article

Salivary Proteomics and Transcriptomics: An Evidence-based Systematic Review

Shivangi Pandey[1], Kuntalika Sarkar[1], Kumar Sougata[1] and Gaurav[2]*

[1]House Surgeons, NSVK Sri Venkateshwara Dental College and Hospital, Bangalore, Karnataka, India

[2]Consultant Oral Physician and Maxillofacial Radiologist, Assistant Professor, Department of Oral Medicine and Maxillofacial Radiology, NSVK Sri Venkateshwara Dental College and Hospital, Bangalore, Karnataka, India

***Corresponding Author:** Gaurav, Consultant Oral Physician and Maxillofacial Radiologist, Assistant Professor, Department of Oral Medicine and Maxillofacial Radiology, NSVK Sri Venkateshwara Dental College and Hospital, Bangalore, Karnataka, India.

Received: February 23, 2021

Published: March 20, 2021

Abstract

Background: Proteins are considered to be the building blocks of our bodies and they are known to harbor a vast array of information, the spectrum of which ranges from the knowledge of physical or pathological conditions to provide us with a pool of biomarkers. Saliva at the same time forms an indispensable part of the oral cavity and with recent researches, it is known to be a huge source of these proteins. As biochemical and molecular technologies are developing each day the gradual intersection of the above two mentioned sources (saliva and proteins) becomes an inevitable pool of knowledge which if carefully studied could throw open future doors of non-invasive diagnostic procedures encompassing multitudes of systemic and metabolic diseases. In this literature review, we have tried and explained the two important biomarkers present in saliva namely proteomes and transcriptomes, and how these are eventually co-related and could serve us to understand the bigger picture of quick and efficient diagnosis.

Aim of the Study: To determine the significance of Salivary Proteomes and Transcriptomes in the field of diagnosis.

Research Question: Are salivary proteomes and transcriptomes significant in supporting early diagnosis?

Materials and Methods: The study sample includes review and research articles based on databases from the Cochrane collaboration as well as few other scientific libraries like Medline and Med-know, having undergone a definite Randomized Control Trial (RCT), to signify the title of the study. Various inclusion and exclusion criterion for Systematic Review decided the final RCT based research articles for the study. (For the given study there were 40 articles which were selected out of which 30 were chosen post having undergone Randomised Control Trial).

Keywords: Randomized Control Trial (RCT); Proteomes; Transcriptomes

Introduction

As the world weaves the cocoon of technological advancement and with the global burden of diseases on the rise it has become but imperative for researches and scientists to come up with new technologies to beat the crisis. In the recent few decades, there has been a growing interest in the scientific community to introduce methods to understand and quantify the study of biomarkers and

allow them to be used as an indicator of disease progression. Long gone are the days of the 1970s when DNA was first sequenced using two-dimensional chromatography, ever since then up till now with the advent of newer technologies such as mass spectrometry and nanotechnology, DNA sequencing and the utilization of genomes in protein biomarker study has improved ten folds. Today it is comparatively easier to accelerate the sequencing process thereby employing the accelerated protein biomarkers in the fields of disease diagnosis and prognosis.

Materials and Methods

Various researches and studies have documented that the use of saliva as a potent diagnostic feature is both specific and sensitive. With this fact in mind, a literature-based systematic review was carried out to fulfill the aim of the study. The study sample includes review and research articles based on databases from the Cochrane collaboration as well as few other scientific libraries like Medline and Med-know, having undergone a definite Randomized Control Trial (RCT), to signify the title of the study. There were 40 articles initially taken and then all they underwent a randomized control trial post which 30 articles were selected for the study.

Results

The saliva secreted is considered to be a protein ultra-filtrate which is either produced *in-situ* or is derived directly from blood thereby can be considered to be an extremely rich source for proteins.

Saliva is an extremely informative bio-fluid that is a readily accessible source and serves as an inexpensive way for harbouring biomarkers.

The knowledge of these biomarkers and their levels in the serum could help us to understand the pathology of the disease thereby allowing us to diagnose it as early as possible therefore increasing chances for improved prognosis.

A proteome is a given set of proteins that is or can be expressed by a genome, cell, tissue, or an organism at a given time under defined conditions and these proteomes are a part of the huge genomic sequencing process, saliva is a clinically proven specific and sensitive proteomic model.

A transcriptome is a collection of all RNA transcripts this includes the coding and the non-coding transcripts in a given population of cells, it is an important step in the final synthesis of proteins and needs to be significantly understood to understand the concepts of protein formation.

Discussion

We have all heard as dentists that the Oral cavity mirrors the health of the body in general, and saliva forms the indispensable part of this oral cavity. It has been known for quite a few years now that saliva is a potent diagnostic agent not only because it serves as a non-invasive method of collection but also because there are several potent biomolecules found in our saliva. Most of the biomolecules circulating in our blood and present in our urine can be easily found in the saliva [1,2]. However, one must remember the quantity of these biomolecules is relatively less ($1/1000^{th}$ than what is present in the blood) [3] thereby calling out for the need for better and more sensitive technologies to be discovered.

The intersection of saliva and diagnosis has to be credited to the advent of molecular and genetic biology. Most of the disease prevalent today can be avoided and their severity lowered if they are diagnosed on time [2,4]. For a diagnostic agent to be good it must be specific, sensitive, and functional and at the same time must meet the characteristics of high throughput, portability all while being cost-effective all while maintaining low cost for the same [5]. It has been known that earlier diagnosis could not only have an impact on the patient's well-being but also improve the overall prognosis as well.

In this paper, we have tried and reviewed the uses and the co-relation of proteomics and transcriptomics and how they are transforming the ideologies of salivary diagnosis in the modern-day world.

Saliva: A multifaceted agent

As we know, human saliva is secreted from three major and several minor salivary glands in majorly three forms i.e. serous, mucous, and mixed [6]. The saliva secreted is considered to be a protein ultra-filtrate which is either produced *in-situ* or is derived directly from blood. Upon stimulation, by the autonomic nervous system [7] saliva is secreted via the salivary ducts into the oral cavity however one must remember that the saliva produced by the serous or the acinar cells are different from the saliva that enters the oral cavity [8]. Once the saliva is produced it travels via the lu-

men of the acinar cells and enters the complex branching systems of ducts. These are the ducts where modifications happen within the saliva they include:- proteolytic cleavage, protein-protein complex formation, and partial de-glycosylation further on entering the oral cavity the fluid gets mixed with several exocrine, endocrine, non-cellular and cellular components, bronchial secretions, oral micro-organisms, remnants of oral wounds, oral debris, desquamated epithelial cells, gingival crevicular fluid and many more additives help the formation of what we know as: Whole Saliva (W.S) [9-11].

With further research one can come across many studies which have time and again mentioned how the protein concentration of saliva is directed under via the parasympathetic control system and their concentration can behavior can be circumstantial depending on a variety of factors such as the cholinergic system, adrenergic system [13-15] or sometimes even under systemic conditions such as xerostomia or radiation therapy can their base constituents change [16]. On average a healthy human being produces-1-2 lts of saliva every day with over 0.5 - 2 mg/ml protein concentration. Now even in a healthy human there exists a discrepancy with the secretion of saliva such as parotid secretions peak around 1 - 2 ml/min between 3 - 5 pm and minimum during 3 - 5 am. Moreover, saliva is also known for its multifaceted function in the oral cavity [16-19].

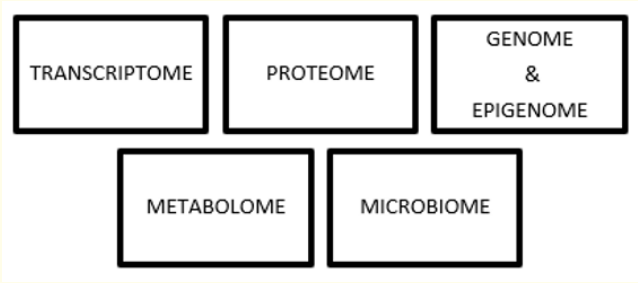
In the light of recent discoveries, scientists have found out about a large number of proteins and biomolecules which could serve as a phenotypic agent for a vast array of diagnostic conditions.

Saliva as a potent diagnostic agent

Most of the systemic complications we know of can be avoidable if they are diagnosed on time thereby guarantying better prognosis, however, most of the time this theory becomes impossible as most of these systemic problems don't show immediate visible specific signs. Moreover, this is the result why the researchers and scientist all across the world are in search of potent biomarkers which can easily study for to better understand the hidden base of most of the systemic problems [20,21]. Now we all know the biomarkers found in our serums if studied: their presence, absence, and rate could serve as a necessary prerequisite to learn more about diseases as specific biomarkers are expressed in specific conditions. Saliva in this case is an extremely informative and readily accessible source for these biomarkers. While on one hand "Whole Saliva" could be used for the process as it is readily collected and contain important serum-derived constituents on the other hand gland-specific saliva can be used to understand the specificity of pathology concerning the major salivary gland [21]. Now a major advantage to the entire process is that sample collection is very simple and comparatively non-expensive, also in most cases, it is non-invasive as well thus reducing the apparent discomfort caused by other invasive procedures for the same making it a viable bio-fluid [23-28].

In recent times there has been a growing appreciation for several proof-principal assays that have continuously helped saliva to utilize its biomarkers for tracing and studying body infections, immune responses, drug reactions and to find out about systemic level problems [26,27]. Moreover, many therapies concerning HIV, cancer, and various other immune-mediated reactions have constantly used saliva for the same. Using human saliva also acts as a huge boon and provides many advantages over the currently available traditional diagnostic approaches. Not only does it serve its purpose of being non-invasive but also it proves to be a stress-free sample collecting technique. It also acts as a source for allowing multiple sampling procedures and proves to be effective in providing reduced risk for pre sampling procedures. It also serves as an effective way of preventing the clinician from contracting any infectious microorganisms such as HPV and HIV and serves for as a perfect diagnostic bio-fluid for the third world countries as it is cost-effective [28].

The biomarkers present in saliva are:



Figure

Proteome and proteomics

The human body harbors several proteins. These proteins are considered to be the building blocks of our system and are of vi-

tal importance. Now a proteome is a given set of proteins that is or can be expressed by a genome, cell, tissue, or an organism at a given time under defined conditions. The etymological origin of the word can be traced back to Marc Wilkins [30,31] who coined it in the year 1994 while being a Ph.D. student at Macquarie University. The word "Proteome" is also a general portmanteau of protein and genomics. Now the concept of proteomes has been applied to a vast number of sources for instances "Tissue Proteome" i.e. the total number of proteins expressed by the given tissue while under an influence of a given set of environment like neural stimulation. Similarly, "Cellular Proteome" is the total number of proteins expressed by a given cell under given circumstances. Now proteome can also be related as a "Complete Proteome" i.e. the total number of proteins expressed by a collection of all the cells in the body or a rough equivalent of "Genome". On the other hand, the word Proteomics refers to the study of these proteins on a large scale. It was first performed on Escherichia Coli. Now there are various methods to study and understand human proteins under proteomics. Generally, proteins can be detected using immunoassays (using antibodies) or mass spectrometry. However, the study is not that simple [29]: The process of proteomics is difficult than its siblings' transcriptomics or genomics because in a given organism the genetic constituent of the genome remains as a constant factor however proteomes differ in different cells and their expression is correlated on basis of situations in which they are made to be expressed. Moreover, it further means that distinct genes are expressed in given cell forms and even the basic set of proteins produced by these cell form need to be assayed [32,33].

In simple words, a cell makes several proteins at any given time unique to the circumstances it is exposed to these circumstances include cell cycle, carcinogenesis, cellular development, and cellular differentiation, as a result, making the process extremely complex. Furthermore, several post-translational modifications make the process extremely difficult.

There are several ways in which proteomic studies are done. Earlier the process was performed under RNA analysis however today it is well known that mRNA does not always translate to form proteins and the amount of protein produced for a given mRNA depends on the gene it is transcribed from and the current state of the cell and its physiology. Moreover, not only does translation from the mRNA cause changes in the proteins produced but also many post-translation modifications occur they range from events such as phosphorylation, glycosylation, methylation, oxidation, nitrosylation, etc.

Human saliva: A proteomic model

The human saliva exhibits as a source for being a potent proteomic model [34,35]. The entire story for salivary biosynthesis begins with the process of gene translation and transcription. The transcription and translation process for salivary proteins occurs in the salivary glands followed by post-translational processes; these include glycosylation, acetylation, phosphorylation and proteolysis. A study was conducted which compared the human salivary proteome to the plasma proteome: as a result, it documented a total of 1939 salivary proteins compiled from a total of 19474 peptide sequence. When compared between the whole and the ductal saliva, 740 of those proteins were identified as a common factor between both. The analyses revealed that there were 3020 proteins out of which 597 were also found to exist in the human saliva. About 99% of the total protein content formed from the variant of plasma proteome was suspected to have 22 abundant proteins in contrast to 20 abundant proteins in the human whole saliva, constituting about 40% of the total protein content. Therefore, this clearly and clinically implies that saliva can be used as a potent diagnostic material to detect biomolecules of clinical sensitivity and specificity with the ease as it is done with blood. However, one must also remember that the whole saliva is highly susceptible to a variety of physiological and biochemical processes, unlike the stable blood plasma. Another important challenge that needs to be expressed is the concentration of the abundant proteins present in the saliva, for example, enzyme amylase is present in saliva at mg/ml concentration however interleukins such as IL-6 and IL-8 (which are clinically extremely important for diagnosis are present in the concentration of pg/ml). The model of the salivary proteome also changes with the advancement of age. In respect to other proteinaceous components, saliva is similar to other body fluids and contains several low molecular weight proteins. These proteins can be grouped into major classes such as histamines, proline-rich peptides, statherins and cystatins. The existence of these different salivary proteins is a result of post-translational modifications. (Therefore for a better understanding of salivary proteins we should understand the protein modification process as a prerequisite to gain insight into the psychological and pathological process-

ing that reflects a person's health and well-being). Human saliva also contains several proteins that aid in maintaining oral homeostasis. In addition to organic and inorganic components saliva also has several disease diagnostic and prognostic factors. For instance, an increased level of salivary thiocyanate has been analyzed to rule out for smokers from non-smokers. Moreover, if one can undergo protein profiling from saliva concerning a disease population, the process can yield valuable clinical parameters to identify various prognostic and therapeutic measures for the same.

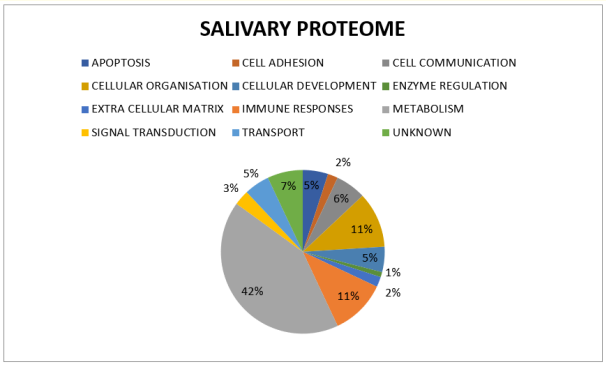


Figure 1: Source: "Proteomic Identification of Salivary Biomarkers of Type (II) Diabetes"- Paturi V. Rao, Ashok P. Reddy, Xinfang Lu, Surendra Dasari, Adiraju Krishnaprasad, Evan Biggs Charles, T. Roberts, Jr., and Srinivasa R. Nagalla.

Transcriptomics and transcriptome

A transcriptome is a collection of all RNA transcripts this includes the coding and the non-coding transcripts in a given population of cells. The term "Transcriptome" is a portmanteau of the words "transcripts" and "genome" and presents with associations during the cellular process of "Transcription" [37]. Transcriptome forms the part of the -home and -omics studies including genome-genomics and proteome-proteomics. The etymological origin of the word transcriptome can be traced back to the formation of cDNA [36] libraries, the first which was published of that of silk moth [38] in the year 1979 following to which the first seminal study to mention the word was published in 1997, the study described the presence of 60633 transcripts expressed in *S. cerevisiae*. With the advent of the broad field of Bioinformatics and subsequently increased in computational power proved to be a cornerstone for the development of the study of transcriptomics. According to the central dogma of molecular theory, the transcriptome comprises firstly protein-coding mRNA transcripts followed by several RNA subtypes with continuous distinct functions.

(It is to also be known that not all types of RNA codes for different proteins the one which doesn't form the non-coding section of the RNAs). If one has to take to account then in the Human Genome about only 5% [40] of all the genes get transcribed into RNA out of which only 1 - 4% form the coding section while the rest forms the non-coding section of RNA. The main aim of the study of the transcriptome is to enlist the given species used for the process of the transcript which include the mRNA, non-coding RNA's and small RNA'S which allow them to determine the transcriptional structures of the genes according to their start sites 5' to 3'. This would also include the splicing patterns, post-transcriptional modifications so that one could quantify the changing expression level of each transcript during given conditions [37,38].

Several factors make the expression of transcriptome difficult to establish:-from the presence of alternative transcription to alternative RNA splicing technique many techniques can capture transcription occurring at a specific point in time. Unlike the genome which is fixed for a given particular cell line, the transcriptome varies per the environment. Moreover, all the mRNA of the cell gets transcribed within the cell itself, the transcriptome provides a perfect reflection of the genes being overexpressed at a given point in time.

The genomic study of stem cells is uniquely important for the research community to understand the nuances of cellular differentiation and carcinogenesis. The processes of gene array data and RNA-sequencing can be used to study the precursor and stem cells to provide independent gene expression data. At the same time, it is a known emerging field to understand the importance of biomarkers and their significance in studying the spectrum of systemic conditions.

One must always remember that the transcriptome is never synthesized de novo because each cell receives a part of their parent transcriptome while the first bough into existence via the process of cell division and maintains its share throughout life. Even

quiescent cells in the bacterial spores have transcriptome; however, the expression of this transcriptome can be switched off.

To simply put it the process of transcription never leads to the formation of transcriptome instead it maintains the transcriptome via replacing the degraded mRNA thereby bringing about the changes in the transcriptome via switching off and on different sets of genes.

Relationship between proteomics and transcriptomics

Although the transcriptome forms merely 4% of the total cell RNA it contains significant information about the coding RNA specifying the composition of the proteome thereby helping to determine the biochemical capacity of the cell. As the information flows from the DNA to RNA via the help of transcription the process doesn't result in unfavorable situations. The major reason behind this is the smooth transition between the DNA and RNA polynucleotides which have similar structural adaptability thereby allowing us to understand how a copy of RNA be made by template-dependent synthesis using base-pairing rules [38,39].

The second phase of genomic expression in which the mRNA [39] molecules of the given transcriptome directly synthesizes proteins is fairly difficult to understand considering the structures of the molecules which are involved. For a long time, the scientists and the researchers were trying to attempt to devise ways and mechanisms via which amino acids could attach to mRNA's, however, this was all in vain because the bonds formed during the experimental attachment deterred the laws of conventional physical chemistry. However, in the year 1957, Francis Crick cut through the process and introduced the idea of the existence of an adapter molecule (this theory later gave rise to the concept of t-RNA). Another problem that arose was the informational dilemma which was further resolved with the understanding of genetic code [39]. There for finally could scientists establish a relationship between the processes of transcriptome and proteome formation specifying how the nucleotide sequence of the mRNA is translated to give rise to the amino acid sequence of the proteins.

Conclusion

Although the use of saliva as a diagnostic agent is promising, most of the diagnostic methods are currently not available for commercial purposes. Moreover, the idea of using saliva as a whole is still in an extremely nascent stage specifically because the amount of protein constituent in the saliva for diagnostic purposes is very minute. Moreover, before we use saliva as a diagnostic fluid we must understand that not only is saliva potent to changes within its biological environment but also the changes are subjective to several unsuited variations such as lifestyle choices, diet modification, age, analytical variations such as sampling error or improper storage and handling conditions. One must also understand that as salivary proteome is extremely sensitive and specific there are several external intrinsic and extrinsic factors care must be taken while documenting the various analytical sources of saliva. With molecular genetics progressing every day and as our knowledge of transcriptomes and proteomes keep on increasing it is quite possible that one day in the near future we will be able to wipe out any source of invasive diagnostic agent for the identification of systemic conditions and this early identification will surely be a boon to several clinicians around the world to finally be able to eradicate several dreadful systemic manifestations in humans.

Bibliography

1. Drake RR., *et al*. "Serum, salivary and tissue proteomics for discovery of biomarkers for head and neck cancers". *Expert Review of Molecular Diagnostics* 5 (2005): 93-100.

2. Pisitkun T., *et al*. "Identification and proteomic profiling of exosomes in human urine". *Proceedings of the National Academy of Sciences of the United States of America* 101 (2004): 13368-13373.

3. Christodoulides N., *et al*. "Application of microchip assay system for the measurement of C-reactive protein in human saliva". *Lab on a Chip* 5 (2005): 261-269.

4. Bouwman FG., *et al*. "2D-electrophoresis and multiplex immunoassay proteomic analysis of different body fluids and cellular components reveal known and novel markers for extended fasting". *BMC Medical Genomics* 4 (2011): 24.

5. Zhang A., *et al*. "Recent and potential developments of biofluid analyses in metabolomics". *Journal of Proteomics* 75.4 (2012): 1079-1088.

6. Herr AE., *et al*. "Microfluidic immunoassays as rapid saliva-based clinical diagnostics". *Proceedings of the National Academy of Sciences of the United States of America* 104.13 (2007): 5268-5273.

7. Humphrey SP and Williamson RT. "A review of saliva: normal composition, flow, and function". *The Journal of Prosthetic Dentistry* 85 (2001): 162-169.

8. Helmerhorst EJ and Oppenheim FG. "Saliva: a dynamic proteome". *Journal of Dental Research* 86 (2007): 680-693.

9. Caporossi L., *et al*. "Saliva as an analytical matrix: state of the art and application for biomonitoring". *Biomarkers* 15 (2010): 475-487.

10. Streckfus CF and Bigler LR. "Saliva as a diagnostic fluid". *Oral Diseases* 8 (2002): 69-76.

11. Dawes C and Jenkins GN. "The effects of different stimuli on the composition of saliva in man". *The Journal of Physiology* 170 (1964): 86-100.

12. Proctor GB., *et al*. "Sympathetic decentralization abolishes increased secretion of immunoglobulin A evoked by parasympathetic stimulation of rat submandibular glands". *Journal of Neuroimmunology* 109 (2000): 147-154.

13. Edwards AV and Titchen DA. "Synergism in the autonomic regulation of parotid secretion of protein in sheep". *The Journal of Physiology* 451 (1992): 1-15.

14. Carpenter GH., *et al*. "Preganglionic parasympathectomy decreases salivary SIgA secretion rates from therat submandibular gland". *Journal of Neuroimmunology* 160 (2005): 4-11.

15. Calvert PA., *et al*. "Autonomic control of submandibular protein secretion in the anaesthetized calf". *Experimental Physiology* 83 (1998): 545-556.

16. Alfredo A., *et al*. "Dysregulated molecular networks in head and neck carcinogenesis". *Oral Oncology* 45 (2009): 324-334.

17. Guajardo-Edwards CFSaC". Biometrics (InTech, USA) (2011).

18. Ferguson DB. "Current diagnostic uses of saliva". *Journal of Dental Research* 66 (1987): 420-424.

19. Dawes C and Jenkins GN. "The effects of different stimuli on the composition of saliva in man*". The Journal of* 170 (1964): 86-100.

20. Hust M., *et al*. "A human scFv antibody generation pipeline for proteome research". *Journal of Biotechnology* 152.4 (2011): 159-170.

21. Perez-Cornejo P., *et al*. "Anoctamin 1 (Tmem16A) Ca2+-activated chloride channel stoichiometrically interactswith an ezrin-radixin-moesin network". *Proceedings of the National Academy of Sciences of the United States of America* 109.26 (2012): 10376-10381.

22. Wang W., *et al*. "Proteomic characterization of transient expression and secretion of a stress-related metalloprotease in high cell density culture of Bacillus megaterium". *Journal of Biotechnology* 126.3 (2006): 313-324.

23. Forsberg L., *et al*. "Pre-fractionation of archival frozen tumours for proteomics applications". *Journal of Biotechnology* 126.4 (2006): 582-586.

24. Will T., *et al*. "Molecular sabotage of plant defense by aphid saliva". *Proceedings of the National Academy of Sciences of the United States of America* 104.25 (2007): 10536-10541.

25. Zauber H., *et al*. "Dynamics of salivary proteins and metabolites during extreme endurance sports - a case study". *Proteomics* 12.13 (2012): 2221-2235.

26. Zhang A., *et al*. "Saliva metabolomics opens door to biomarker discovery, disease diagnosis, and treatment". *Biotechnology and Applied Biochemistry* (2012).

27. Wu Z., *et al*. "Quantitative chemical proteomics reveals new potential drug targets in head and neck cancer". *Molecular and Cellular Proteomics* 10.12 (2011): M111.011635.

28. Nagler R., *et al*. "Heparanase up-regulation in tongue cancer: tissue and saliva analysis". *Cancer* 110.12 (2007): 2732-2739.

29. Epstein JB., *et al*. "The correlation between epidermal growth factor levels in saliva and the severity of oral mucositis during oropharyngeal radiation therapy". *Cancer* 89.11 (2000): 2258-2265.

30. Anderson NL and Anderson NG Anderson. "Proteome and proteomics: new technologies, new concepts, and new words". *Electrophoresis* 19.11 (1998): 1853-1861.

31. Wilkins Marc. "Proteomics data mining". *Expert Review of Proteomics* (2009).

32. Wasinger VC., *et al*. "Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*". *Electrophoresis* 16.1 (1995): 1090-1094.

33. Vikas Dhingraa., *et al*. "New frontiers in proteomics research: A perspective". *International Journal of Pharmaceutics* 299.1-2 (2005): 1-18.

34. L Schulz., *et al.* "Saliva proteome research: current status and future outlook". *Critical Reviews in Biotechnology* 33.3 (2013): 246-259.

35. Aihua Zhang., *et al*. "Salivary proteomics in biomedical research". *Clinica Chimica Acta* 415 (2013): 261-265.

36. Croy Ron. "Molecular Genetics II - Genetic Engineering Course (Supplementary notes)". Durham University durham.ac.uk; 20 April 1998. Archived from the original on 24 August 2002 (2015).

37. Wang Zhong., *et al*. "RNA-Seq: a revolutionary tool for transcriptomics". *Nature Reviews Genetics* 10.1 (2009): 57-63.

38. GK Sim., et al. "Use of a cDNA library for studies on evolution and developmental expression of the chorion multigene families". *Cell* 18.4 (1979): 1303-1316.

39. Brown TA. "Genomes". 2nd edition. Oxford: Wiley-Liss; Chapter 3, Transcriptomes and Proteomes (2002).

40. C Frith Martin., *et al*. "Genomics: The amazing complexity of the human transcriptome". *European Journal of Human Genetics* 13.8 (2005): 894-897.