Chapter

The Salivary Secretome

Luís Perpétuo, Rita Ferreira, Sofia Guedes, Francisco Amado and Rui Vitorino

Abstract

Recently, proteomics has emerged as an important tool for understanding biological systems, protein–protein interactions, and networks that ultimately lead to a deeper understanding of the underlying mechanisms of certain diseases. More recently, the study of secretomes, a type of proteomics, has also been highlighted as a potential next step in the field of diagnosis/prognosis. The secretome is the set of proteins expressed by an organism and secreted into the extracellular space, comprising 13–20% of all proteins. Since almost all, if not all, organs produce secretomes, this means that it is possible to study secretomes and trace these proteins back to their origin, supporting the idea that this could indeed be very important in diagnosing certain diseases. This is often combined with techniques such as mass spectrometry to measure the secretome of, for example, a particular tissue, and bioinformatics tools and databases to give us an idea of what to expect (prediction). In this paper, we will give a general overview of this world, but with a focus on the new bioinformatics tools and databases, their advantages and disadvantages, as well as a deeper look at isolation systems for proteomes, specifically salivary secretomes. Indeed, the salivary secretome represents a valuable new tool capable of providing insights into immunopathology and potentially aiding in diagnostics. Furthermore, we will explore applications of these methods and give an idea of what the future holds for such promising techniques: Salivary secretome in conjunction with bioinformatics tools/databases in the diagnosis of diseases (such as diabetes, Sjogren's syndrome, and cardiovascular disease).

Keywords: Saliva, bioinformatic tools, prediction

1. Introduction

Salivary plasma is also known as ultrafiltrate of biological fluid. Nearly 1,000 different proteins and 19,000 unique peptide sequences have been found in saliva. Whole saliva (WMS) is a combination of various secretions produced by major and minor salivary glands, gingiva cervical fluid (GCF), mucosal transmission, oral wound serum and blood vessels [1].

Saliva is mainly composed of three main pairs of salivary glands, namely the parotid gland, the sublingual gland, and the lingual gland. Saliva is composed of 99.5% water, 0.3% protein and 0.2% trace elements. The concentration of proteins and peptides in saliva is very important for the maintenance of oral health and homeostasis. The increased frequency and severity of oral diseases are often related

to changes in salivary protein content [2, 3]. Salivary proteins facilitate food perception and digestion, maintain the integrity of mineralized tooth and oral epithelial surfaces, shield the oral digestive tract from environmental hazards and invading pathogens, and protect oral tissues from fungal or viral infections [4, 5]. Moreover, it is suggested that the origins of salivary proteins can be analyzed throughout mixed saliva and that post-transcriptional modifications may play a key role in understanding secretome network pathways [6, 7]. Moreover, Feizi et al., 2020 [7] found that the study of secretion pathways (translocation, folding, trafficking and glycosylation) is relevant to know tissue-specific secretion pathways and tissue-specific secretomes that would facilitate the elaboration of a link between proteins and diseases.

Salivary gland secretome represents a valuable new tool to measure many local soluble mediators to gain future insight into immunopathology and potentially aid in diagnosis [8]. This method could be of use to identify therapeutic targets and develop markers for stratification, prognosis and treatment response in patients [9, 10].

Biomarkers are defined as biological molecules found in blood, saliva, and other body fluids, as symptoms of normal or abnormal processes, or symptoms of conditions or diseases. Few studies have demonstrated the relationship between serum and saliva levels of clinically used biomarkers. Nevertheless, human saliva has attracted attention as a liquid biopsy for the diagnosis of oral diseases as a potential target. Salivary biomarkers are used for evaluation, prediction and diagnosis of various diseases. They can be collected rapidly in a non-invasive, natural and painless way. Salivary research can help identify biomarkers associated with health and disease conditions [1, 11, 12]. Collection and storage of saliva is also simple, relatively cheap and low risk for patients and healthcare professionals [3, 13, 14]. Salivary proteomics is a promising tool as proteomic molecules control the direct antibacterial action of microorganisms in the oral cavity, but has limitations including non-applicability in the driest patients and technical challenges such as the degradation of cytokines by salivary enzymes [9]. Optimization of existing histology protocols to determine salivary gland inflammation will help improve the diagnosis of various diseases [9, 10].

2. Definition

Under both normal and pathological conditions, cells secrete a variety of proteins into the extracellular space via classical and non-classical secretory pathways [5]. The majority of these proteins represent the pathophysiology of the cell from which they are secreted. Recently, although more than 92% of protein-coding genes have been mapped by the Human Proteome Map Project, a large number of these proteins that constitute the cell's secretome is still unknown [4].

Secreted proteins or the secretome may be accessible in body fluids and are therefore considered as potential biomarkers to distinguish between healthy and diseased individuals [8]. To facilitate biomarker discovery and further assist clinicians and scientists working in these areas, we compiled and cataloged secreted proteins from the human proteome using an integrated bioinformatics approach [9, 10]. In this study, it was found that nearly 14% of the human proteome is likely secreted via classical and non-classical secretion pathways. Of these, ~38% of these secreted proteins were found in extracellular vesicles including exosomes and shedding microvesicles.

Of these secreted proteins, 94% were detected in human body fluids including blood, plasma, serum, saliva, semen, tears, and urine. We hypothesize that this list of secreted proteins with high confidence could serve as a compendium of biomarker

candidates. In addition, the catalog could provide functional insights into understanding the molecular mechanisms involved in various physiological and pathophysiological conditions.

The highly elevated inflammatory mediators in the secretions of patients with complex diseases such as primary Sjögren's syndrome (PSS) are related to clinical parameters.

2.1 Secreted proteins

A secreted protein can be defined as a protein that is actively transported outside the cell. In humans, cells such as endocrine cells and B lymphocytes are specialized in the secretion of proteins, most cells secrete proteins in different. Not only is this a rich source of new treatments and drug targets, but most of the blood diagnostic tests used in the clinic target secreted proteins, underscoring the importance of these proteins to medicine and biology. These include pancreatic enzymes (PRSS1, CELA3A, AMY2A) and other digestive enzymes expressed in the salivary glands (PRR4, STATH, ZG16B) or stomach (PGA3, PGA4). The liver is one of the most important secretory organs and produces high amounts of plasma proteins such as albumin, fibrinogen and transferrin. Another group of highly secreted proteins belongs to the diphenhydramine family and is secreted by glandular cells in the epididymis (DEFB118, DEFB106A and DEFB129).

2.2 Membrane proteins

Membrane proteins are one of the largest and most important classes of proteins. Membrane proteins are associated with cell membranes or organs in cells and can be classified as peripheral or integrated. Peripheral membrane proteins bind to the membrane by binding to the peripheral region of the membrane or by integrated membrane proteins, but cannot completely penetrate the membrane. Integrated membrane proteins have a hydrophobic α-helical or β-barrel structure so that they can be distributed throughout the lipid molecule and linked by the outer loop region of the membrane. The α-helical integral membrane protein is the main type of membrane protein and is found in all types of biological membranes. This explains why their key roles as transporters and receptors currently account for about 57% of approved drug targets, as they are of great importance to the pharmaceutical industry. Many important receptors and cell surface molecules are found in the list of human cell differentiation molecules (CD markers). The G protein synthesis receptor (GPCR) comprises seven transmembrane fragments (DM) and contains 775 human protein-coding genes, making it the largest membrane protein target.

2.3 Classification of the human proteome

Despite the availability of the human salivary proteome, the origin of individual proteins remains unclear. So far, more than 3000 proteins have been identified in various studies, and with new tools and methods, more will be identified [4].

Meinken et al., 2015 [15] analyzed the subcellular location of the protein using MetazSecKB. The subcellular location of the protein is an important factor that determines the function of the protein molecule in the organism. MetazSecKB is a knowledge base for subcellular proteins developed specifically for metazoans (i.e., humans and animals). More than 4 million protein sequence data entries have been retrieved from UniProt, including 121 complete proteins. The location of protein

subgroups, including secretion and 15 subgroup sites, are assigned based on selected test evidence or predictions using 7 computational tools [15].

Various identifiers, gene names, keywords and types can be used to search and download protein or subcellular protein data. Support BLAST search and community annotation of sublocalizations. Our preliminary analysis shows that protein levels, secretome levels, and other subcellular protein levels vary widely among different animal species. Confidentiality levels range from 3–22% (mean 8%) in Metazzoa species [15].

Approximately 21–43% (mean 31%) in cytoplasm, 20–37% (mean 30%) in embryo, 3–19% (mean 12%) plasma membrane proteins and 3–9% (mean 6%) in mitochondria. The authors also compared protein families in different animal species [15]. Genetic oncology of human secreted proteins and field analysis of protein families show that these proteins play an important role in the development of human structure, signal transduction and regulation of many biological processes in the immune system [15].

The combination of the results of membrane protein and secretion analysis draws the distribution map of potential membrane proteins and secreted proteins in human membrane proteins. Three types are used to annotate the protein isoforms of all human genes: (i) secreted type, (ii) membrane type, and (iii) endogenous type (i.e., proteins with no predictable SP/TM properties). Note that proteins classified as membranous may be localized in the endoplasmic reticulum or in the inner membrane, such as the colon. Each human protein-coding gene is classified as having all isoforms encoding protein isoforms belonging to one or two or three types of these groups. The results showed that at least 36% of the predicted human genes contain membrane-disseminated or secreted protein isoforms.

2.4 The plasma proteome

Plasma is the transparent, liquid part of blood that is formed when white blood cells, red blood cells, and platelets are removed. It consists of small substances such as water (90%), protein (7–8%), salt, gas, and nutrients. Plasma proteins contain up to 90% of the ten most abundant proteins, including albumin, fibrinogen, which is involved in blood clotting, and immunoglobulin, which is mainly involved in immune processes. One of the most important functions of plasma is to transport essential compounds to different parts of the body, regulate osmotic pressure and fluid exchange in all tissues, and play an important role in immune system function. Most cells in the body interact with plasma directly or indirectly through other fluids. Therefore, analysis of plasma proteins can provide important information about the patient's health status.

The dynamic range of plasma protein between the high albumin (ALB) concentration is more than 10 orders of magnitude. It has an extraordinary dynamic range. It can serve as a transporter and helps maintain the osmotic pressure of emulsification. Rare proteins containing interleukins are found in tissues. Although many proteins in the plasma proteome pass through the secretory pathway, there is another type of tissue-secreted proteins that are found in cells but can be released into plasma due to cell death or damage. There is also an interesting class of proteins that do not enter the ER/Golgi pathway by non-classical secretion and include cytokines such as interleukin 10 (IL10) and mitogens such as fibroblast growth factor 2 (FGF2).

2.5 The secretory pathway

In the secretory pathway, the signal sequence protein travels from the endoplasmic reticulum (ER) through the vesicles of the ER to the cell surface. The signal

sequence that drives protein secretion is called a signal peptide. It is a short hydrophobic N-terminus that is inserted into the ER membrane and separated from the protein. In most cases, the N-terminal transmembrane (TM) acts as part of the signal line. The ER signal sequence is recognized by the chaperone protein, which guides the ribosome to the approximate ER where transfer of the protein sequence takes place in a protein complex called the translocon. The membrane protein is transferred by the translocation protein to the lipid player of the ER membrane, and the secreted protein is transferred into the lumen of ER. Proteins that pass the quality control of ER are transported by vesicles to the Golgi, where they are further modified in important processes such as glycosylation and phosphorylation. The Golgi is also responsible for sorting proteins for transport to their final destination. These proteins are usually plasma membranes, lysosomes or cell secretions.

3. Systems for isolation

With the use of novel advanced technologies, many oral and systemic diseases can be treated early with non-invasive, easy to follow, time-saving and personalized solutions, further enhancing the potential of salivary secretome [16].

3.1 Common methods used to identify salivary secretome

Salivary secretome mainly includes proteins, metabolites, genes, microorganisms and immune system. Various methods are used to analyze molecules to study and verify biomarkers. Proteins can be used to diagnose diabetes, periodontitis, dental caries and AIDS [16, 17]. However, salivary transcriptome and genes include mRNA and DNA. Genetic chip sequencing, DNA hybridization, qPCR and gel electrophoresis help in identifying various diseases. On the other hand, metabolic research requires and uses gas chromatography–mass spectrometry, nuclear magnetic resonance spectroscopy and high-performance liquid chromatography. These methods can be used to diagnose diabetes, lung cancer, pancreatic cancer, breast cancer and Sjogren's syndrome [17, 18].

3.2 Saliva biomarker-based platforms

The analysis system is based on different technologies used to detect biomarkers in saliva. Single and multiple systems (e.g. MEMS, ORI, chromatographic test strips and multiple salivary glands (US)) are only used to detect proteins and whole proteins and nucleic acids (e.g. IL -8, MMP-8, α-amylase, e.g. IV, HCV) up to 1 minute [19], and there is a short time limit. Therefore, these technologies have reduced the aggressive behavior to a higher level [19].

3.3 Novel isolation techniques

Biosensor is a biological analyzer that can replicate any biological material. The biosensor works by biometrically identifying specific components and is designed for target analysis. They remain selective and sensitive to the presence of other interfering compounds. In the medical field, the application of biosensors is growing rapidly [20, 21]. They can detect antibodies/antigens, nucleic acids, cell structures or enzymes. The transducers can be electrochemical, thermometric, optical, piezoelectric or magnetic.

3.4 Automated mass spectrometry-based approach

Using liquid chromatography–tandem mass spectrometry (LC -MS/MS) to identify novel targets in specially prepared cell secretions, Wetie et al., 2013 [22] have developed an automated, simple and effective strategy. In addition, the supporting role of mass spectrometry (MS) in the functional evaluation of the identified secreted targets is investigated [22].

Simplicity is achieved by culturing cells in serum-free medium, which eliminates the need to remove large amounts of serum protein while minimizing unsightly matrix effects. Once these factors have been determined, their verification and nature is followed. In addition, this method can lead to the identification of abnormally secreted, spilled or exaggerated proteins in response to stimuli [22].

3.5 BioMEMS

The lab-on-a-chip system uses small and easy-to-build BioMEMS devices to detect biological and chemical agents. They are based on micro/nano scale fabrication systems and help to improve the sensitivity of sensor results. It has unlabeled detection technology including microconverter, surface plasmon oscillation and organic field effect transistors [23, 24].

Biological Micro-Electro-Mechanical Systems (MEMS) can be used in many applications, such as drug delivery, heart MEMS, hearing aids, insulin microbumps, endoscopic lens agents, and retinal prostheses for monitoring patients with heart disease [25].

3.6 Fluorescent biosensors

Fluorescent biosensors can be used in cancer, drug development, arthritis, cardiovascular and viral infections, chronic myeloid leukemia, etc. Efficient screening methods, applications of fluorescence studies in gene expression, protein location in cell cycle, cell apoptosis, signal transduction and transcription [26]. Nanomaterials and nanoproducts for biosensors offer opportunities for the next generation of biosensing. They can be widely used for monitoring, diagnosis, control and analysis [27].

3.7 Electrical field-induced release and measurement (EFIRM)

The liquid biopsy technique called EFIRM uses electrochemical methods to promote hybridization of nucleic acids [28]. This method enables precise detection of RNA and protein biomarker targets on exosomes. EFIRM can analyze mutation status within one hour without extracting DNA. It can detect cancer in oral cavity cancer, non-squamous lung cancer and epidermal growth factor (EGFR) mutations [29].

3.8 Microfluidics

Microfluidic applications work with integrated micromachining and specific physical and chemical properties. Originally, silicone, mineral glass and ceramics were used. These materials were replaced by soft and hard thermostatic and thermoplastic materials or biodegradable hydrogel materials.

The paper analytical device (μPADs) was first developed by Whiteside. Paper is microscopic and hydrophilic, so it provides the basis for the formation of microscopic channels. They are used to diagnose urine metabolism, blood glucose, pH, liver function and infectious agents [30].

4. Interaction with other systems

Saliva is a potential diagnostic tool that can provide a simple diagnostic method. The presence of salivary biomarkers can aid in early diagnosis. Saliva has the potential to revolutionize next-generation molecular testing. It can diagnose oral cancer, dental inflammation and periodontitis [31]. To date, many salivary biomarkers have been proposed for the diagnosis of oral cancer. Conventional medical standards are not sufficient to easily determine the location of active disease or to easily measure the progress of future disease. Genetic testing offers the most effective way to prevent dental disease in the long term [31–33].

In addition, several research groups have reported the use of whole saliva or glandular saliva for mass spectrometry-based proteomics research. The extensive enumeration of salivary proteins is done by combining the previous LC -MS/MS -based saliva research data with our research data [34].

Using the bioinformatics tools mentioned above, a possible analysis of gene ontology classification and their secretion was performed. Comparing with the latest human salivary proteins synthesized from oral cancer tissues expressing different proteins, we found proteins associated with oral cancer. The protein peptides of these proteinsor the most observed peptides were selected from the Global Protein Machine Database (GPMDP) [35].

5. Bioinformatics tools for secretome prediction

Proteins can flow from blood to salivary glands by active transport, passive diffusion, or ultrafiltration, and then some of them are released into saliva, so if accurately identified, they can be used as biomarkers of disease [36]. Researchers have developed a series of novel computational and biological communication tools to predict salivary biomarkers [37].

The basis of the prediction is a set of physicochemical and hierarchical features found between human proteins that can pass from blood to saliva and proteins that are not present in saliva [36, 37]. In 2013, Wang et al., [38] predicted human salivary proteins from blood and evaluated their use in identifying diagnostic biomarkers. This predictive capability can be used to predict potential biomarker proteins for specific human diseases, information about various exposed proteins in diseased and healthy control tissues, and the prognostic potential of proteins secreted in blood. This enables the use of antibody-based technology to target effective biomarkers in saliva. They used this comprehensive data to predict that 31 candidate biomarker proteins in saliva could be used in breast cancer [38].

5.1 Bioinformatic tools

Continued method development supports the comprehensive identification and quantification of secreted proteins at specific cellular levels. The role of secretory factors in regulating important signaling events has been discussed, and a connectivity diagram has been constructed to describe differential secretory expression and dynamic changes [39].

Bioinformatics has become a bridge between confidential data and computer tasks to manage, mine and retrieve information. Based on this information, predictions can be made to help clarify the physiological state of a particular organism and determine

the specific dysfunction at the stage of disease. The major challenge in data analysis lies within the integration of biological information from different sources. Database enhancements and software improvements can greatly increase the practicality and reliability of confidential investigations [39]. Reliable data interpretation is essential for the formation and exploration of relevant disease biomarker proposals as well as the discovery of new drug targets. Using genetic oncology (GO), it is possible to collect basic information about secretome proteins. GO analysis can determine how the identified components relate to specific functions or processes and whether a particular type of protein is found in secretions. In addition, in a statistical framework, the method GO can determine whether there is an obvious GO period [40]. In the database GO, molecular functions are defined as the biochemical functions of gene products. The biological pathway refers to the integration of the biochemical properties of proteins. Pathway analysis is an important step to properly understand the uniqueness and function of secrets. Methods have been developed to assess whether protein packaging is present in the target phenotype. Using different tools, according to different group rules, can provide different and sometimes complementary information. The results are strongly influenced by the criteria chosen to define the target protein and the reference list [40, 41].

Ingenuity Path Analysis (IPA) and Meta Core (Genico) are commercial software for visualizing high performance data in a biological network environment. STRING is a free database of known and predicted protein interactions, including direct (physical) and indirect (functional) associations from various sources [42]. In this way, the network of protein–protein interaction, metabolism or genetic regulation can be reconstructed based on prior knowledge and the biological network can be reconstructed. Determined by the interaction between its components [42, 43].

Pathway analysis tools become very popular and can interpret omics data quickly. To date, IPA has been highly cited in the field of proteomics, having been used in 121 publications. The software is designed to interpret large genetic datasets, but it can also be used to illustrate the biological implications of complex proteomics datasets [43].

Data sharing presents a new challenge for modern proteomics. The first obstacle to data sharing is the data format. Each MS tool generates a file from the source data in a proprietary format. The HUPO-PSI standard has been accepted by vendors and public web-based resource providers. The proteomics community has developed guidelines to facilitate storage and open access to proteomics data in a central public repository [44–46].

PeptideAtlas, PRIDE and Trench have been developed to share data among the entire proteomics research community [47–49]. Recently, the Proteome Change Alliance has established a place where MS proteomics data can be submitted to the existing major proteomics databases. The purpose is to facilitate data transfer between them to achieve the best data transfer and to create a global accession number for all participating databases. This information will be available to all MS/MS researchers in the UK who wish to use it for their research [47–49].

Using the innovative majority voting methods, Rehman et al., 2020, analyzed transcriptome data from 5 cancer types and more than 3000 samples to measure the relative difference in gene expression of secretory proteins compared to normal tissue in the vicinity. A comprehensive, in-depth data mining analysis reveals that among several cancer types, a continuous group of uncontrolled secretory protein subtypes is concentrated in hematopoietic cell lines. Genes associated with hematopoietic cell lineages are often reduced during the continuous development of cancer, and high exposure levels are associated with good prognosis for patients [50].

Moreover, they suggest that cancer cells suppress the underlying mechanism of hematopoietic cell lineage signaling by reducing the expression of immune-related genes. The data identify potential biomarkers for cancer immunotherapy. It can be concluded that this method is applicable to define other cancers and highlight specific targets for treatment and diagnosis [50].

5.2 BONCAT and pulsed SILAC

Despite the increasing interest in secretomes associated with paracrine/autocrine mechanisms, mass spectrometric cell studies have been performed using serum-free media (SFM). On the other hand, the use of serum culture medium (SCM) is not necessarily recommended because secretions obtained with SCM are easily contaminated with fetal serum proteins (FBS) [51].

Shin et al., 2019, [51] used biological non-designated amino acid tags (BONCAT) and pulsed SILAC (pSILAC) to analyze the different secreted proteins between SFM and SCM. Mesenchymal stem cells are derived from human cancer cells U87MG and human Wharton's Jelly (hWJ-MSCs) [51]. In most cases, the biological communication equipment predicts that the protein is secreted when the protein secretion level in SCM is higher than in SFM. In HWJ-MSC, the amount of protein secreted in SCM within 24 hours, even considering different cell proliferation rates, is greater than SFM [51]. The highly secreted HWJ-MSC protein in SCM contains many positive markers of angiogenesis, neurogenesis and osteogenesis, as well as MSc-paracrine factors involved in upstream regulators of cell proliferation. This result indicates that secretome analysis should be processed in SCM to promote cell proliferation and secretion [51].

Another computational method was evaluated by Min, 2010 [52] to understand the prediction accuracy of signal B, phobia, target B and wolf sport, which can be used alone or in combination with DMHMM and PS scanning. Prediction accuracy is represented by Mathews Correlation Coefficient (MCC). Tools for predicting proteins secreted in different eukaryotic kingdoms show different advantages. Using his own tools, the author found Wolfsport for fungi (73.1%), Phobius for animals (82.8%), Signal B for plants (55.4%) and Phobius for proteases (42.1%) [52].

The use of TMHMM significantly improves the prediction accuracy of all datasets. According to the measured accuracy, it is recommended to use the following methods to make secret predictions for different eukaryotes: signal P/DMHM/wolfport/Phobius/PS scan for fungi (83.4%), Phobius/wolfb/animal/PS -86A Phobius/target P/PS scan (73.2%), combined with all tools for protists (52.8%) [52].

Free interactive resources are provided within the portal Human Protein Atlas (www.proteinatlas.org) and analyzed by Uhlén et al., 2015 [53]. The portal offers the possibility to explore tissue-expressed proteins in tissues and organs and to analyze tissue profiles for specific protein classes. A large list of proteins expressed at high levels in different tissues was compiled to localize the protein in the subunits of each tissue and organ and provide a spatial environment down to the level of individual cells [53].

6. Applications

Saliva is a complex fluid containing various enzymes, electrolytes, proteins, nucleic acids, antibacterial components, hormones, cytokines, and antibodies.

Its composition almost reflects the overall state of physical health and disease. It can become a diagnostic tool for many diseases. The submandibular saliva of patients with cystic fibrosis contains 66% more lipids per 100 milliliters of saliva than healthy substances. Salivary fatty acid profile can be used as a good indicator for early detection of heart disease. Dietary fat intake influences the increased arachidonic acid production associated with lung inflammation and heart disease.

Under normal and pathological conditions, cells secrete various types of proteins into the extracellular space via classical and nonclassical secretory pathways. Most of these proteins represent cell secretion pathologies. Recently, Human Protein Atlas Project has localized more than 92% of protein-coding genes, but the number of proteins secreted by cells is still difficult to determine [54]. Secreted proteins or secretions can enter body fluids and are therefore considered as potential biomarkers to distinguish healthy and diseased individuals. To facilitate the discovery of biomarkers and to further assist physicians and scientists working in this field, Keerthikumar et al., 2016, [54] used integrated bioinformatics methods to compile and list the secreted proteins in humans.

In this study, it was found that about 14% of human proteins can be secreted through classical and non-classical secretion pathways. Among them, about 38% of secreted proteins are in extracellular cells, including exosomes and excretory microorganisms. Of these secreted proteins, 94% are present in human body fluids, including blood, plasma, serum, saliva, semen, tears, and urine [54]. The author hypothesizes that this list of secreted proteins can serve as a set of candidate biomarkers with high confidence. They can provide functional insights to understand the molecular mechanisms associated with various physiological and pathophysiological states of cells [54].

Chen et al., 2019, [55] found that secretory proteins are widely expressed in various tissues and body fluids, and a large proportion of them are expressed in a tissue-specific manner. In addition, there are 14 cancer-related secretory proteins. Their expression levels are significantly correlated with survival rates of patients with eight different tumors, which may be potential prognostic biomarkers [55]. Surprisingly, of the 6,943 secretory proteins, 89.21% (2,927 novel secretory proteins) have known protein domains [55]. The authors enriched these novel secretory proteins mainly by known domains related to immunity (such as immunoglobulin V set and C1 set domains). Their comprehensive novel secretory proteins and features provide insight into human confidentiality and are valuable resources for future research [55].

In Sjogren's syndrome, salivary flow is impaired due to tubular changes caused by lymphocyte infiltration and salivary gland fibrosis, and the patient suffers from toothache, infectious dysphagia, and other oral complaints. The blood lipid level of Sjogren's patients is twice that of normal healthy people, and the antibody level is high. Patients with Sjogren's syndrome or radiation therapy for head and neck cancer have severe dry mouth, which greatly affects their oral health and quality of life. Since there is no clinically proven treatment, clinical management of xerostomia is limited to preventive treatment. Previous research has shown that mesenchymal stem cells (MMSC) derived from mouse bone marrow differ from salivary progenitors when grown together with mouse salivary epithelial cells [56]. Restrictive transcription factors in co-grown MMSCs are identified with amylase (AMY1), muscarinic 3 receptors (M3R), aquaporin 5 (AQP5), tubular morphological changes, and acinar cell

marker expression [56]. This cell marker is called cytokeratin 19 (CK19). Mona et al., 2020, investigated inducible molecules in a conditioned medium that can trigger MMSC replication and integrated mass spectrometry and systems biology by high

Bioinformatic tool/database	Features	Accessibility	Website
BIOCARTA	Protein-pathway association	B	https://maayanlab-cloud/Harmonizome/dataset/Biocarta+Pathways
BLAST	Sequence similarities comparison	C	https://blast.ncbi.nlm.nih.gov/Blast.cgi
DisGeNET	Gene and variability associated with disease	B	https://www.disgenet.org/
Gene Ontology	Computational model of biological systems (genes)	A	http://geneontology.org/
GPM	Proteomics data analysis, reuse and validation for biological and biomedical research	C	https://www.thegpm.org/
IPA	Genomic and clinical knowledge	A	https://digitalinsights.qiagen.com/
KEGG	Database of high-level functions on biological systems	C	https://www.genome.jp/kegg/
Meta Core	Omics analysis	B	https://portal.genego.com/
MetazSecKBGene	Knowledgebase for human/animal secretomes	B	http://proteomics.ysu.edu/secretomes/animal/index.php
Panther	Biological pathways	B	http://www.pantherdb.org/
PathwayStudio	Network between omics and biological processes	A	https://www.pathwaystudio.com/
PeptideAtlas	MS/MS peptide data	B	http://www.peptideatlas.org/
Philius	Signal peptide prediction	A	http://www.yeastrc.org/philius/pages/philius/runPhilius.jsp
Phobius	Signal peptide prediction	B	http://phobius.sbc-su.se/
PRIDE	Proteomics datasets	B	http://www.ebi-ac.uk/pride/
SecretomeP	Non-classical secretion prediction	A	http://www.cbs.dtu.dk/services/SecretomeP/
SignalP	Signal peptide prediction	A	http://www.cbs.dtu.dk/services/SignalP/
STRING	Protein–protein interaction networks	A	https://string-db.org/
TMHMM	Transmembrane helix prediction	A	http://www.cbs.dtu.dk/services/TMHMM/
UniProt	Protein database	B	http://www.uniprot.org/
Wolfpsort	Subcellular localization prediction	B	https://wolfpsort.hgc.ip/

A - easy to use and accessible, B - accessible but hard to use, C - hard to access and use.

Table 1.
Bioinformatics tools and databases that predict secreted proteins.

performance liquid chromatography. Based on their key roles in embryonic development and salivary gland growth, our method identified ten differentially expressed proteins [56]. In addition, systems biology analysis revealed six candidate proteins, namely cysteine-rich insulin-like growth factor binding protein 7 (IgFPP7), pro-angiogenic stimulant 61 (CYR61), acrin (AGRN), laminin, beta 2 (LAMP2), folistatin 1 (FSDL1) and fibronectin 1 (FN1), all of which could potentially contribute to the propagation of MMSC during co-cultivation [56].

Human salivary secretome plays a diagnostic role in the diagnosis of heart disease. Diabetes is another common disease that is rapidly developing worldwide. Due to its non-invasiveness, cheap and simple saliva samples are attractive as diagnostic fluids for diabetes analysis. The announced study concluded that various biomarkers are used in the early stages to diagnose diabetes. Compared to serum of diabetic and diabetic patients, salivary glucose, amylase, calcium, phosphorus and calcium levels show significant changes.

Salivary secretome test requires proper identification and verification of biomarkers. Diagnosis and biomarkers are measurable parameters that can interact physiologically and biochemically at the molecular or cellular level and always serve as normal indicators. The pathology and intervention behavior of the human body can be identified using biomarkers present in salivary secretome. Biomarkers include many categories, such as protein, DNA, RNA, metabolism and microorganisms, so they are all used together (**Table 1**).

7. Future perspectives

With our current results, we note that although the secretome has gained attention and has been highlighted in recent studies, it is still of interest to explore this topic more deeply. In future studies, we propose to go beyond the usual protein profiling and perform network studies to find links between proteins from salivary secretome in direct and indirect ways. In addition, studies have begun to evaluate the role that transcriptional and post-transcriptional modifications of proteins have in informing their origin [6, 7]. This will be relevant for establishing links between salivary protein levels and disease prognosis/diagnosis.

In **Figure 1**, we see a flowchart of the information presented in this paper. Starting from a secretome: how do we detect it (detection approach), how do we identify it (protein identification), how do we verify information about these proteins (record repositories and inventory of the secretome). This line of work leads to the analysis of the secretome, but makes up only part of the pipeline. Prediction tools are also of great use (genomic datasets enable the existence of tools to predict secretory proteins). Several of the bioinformatics tools discussed previously can be used to perform secretome analysis, where we can limit the investigation to protein profiling, but also go beyond that to investigate signaling pathways, networks (protein–protein, gene-protein, protein-disease), and determine useful disease biomarkers. All of this information culminates in getting closer to the biological significance of certain proteins and interactions under certain circumstances. **Figure 1** shows several examples of biomarkers mentioned in this review and narrows that down to applications for saliva testing, namely in areas such as prognosis and diagnosis. In summary, **Figure 1** represents the pipeline and workflow of secretome studies.

DOI: http://dx.doi.org/10.5772/intechopen.98278

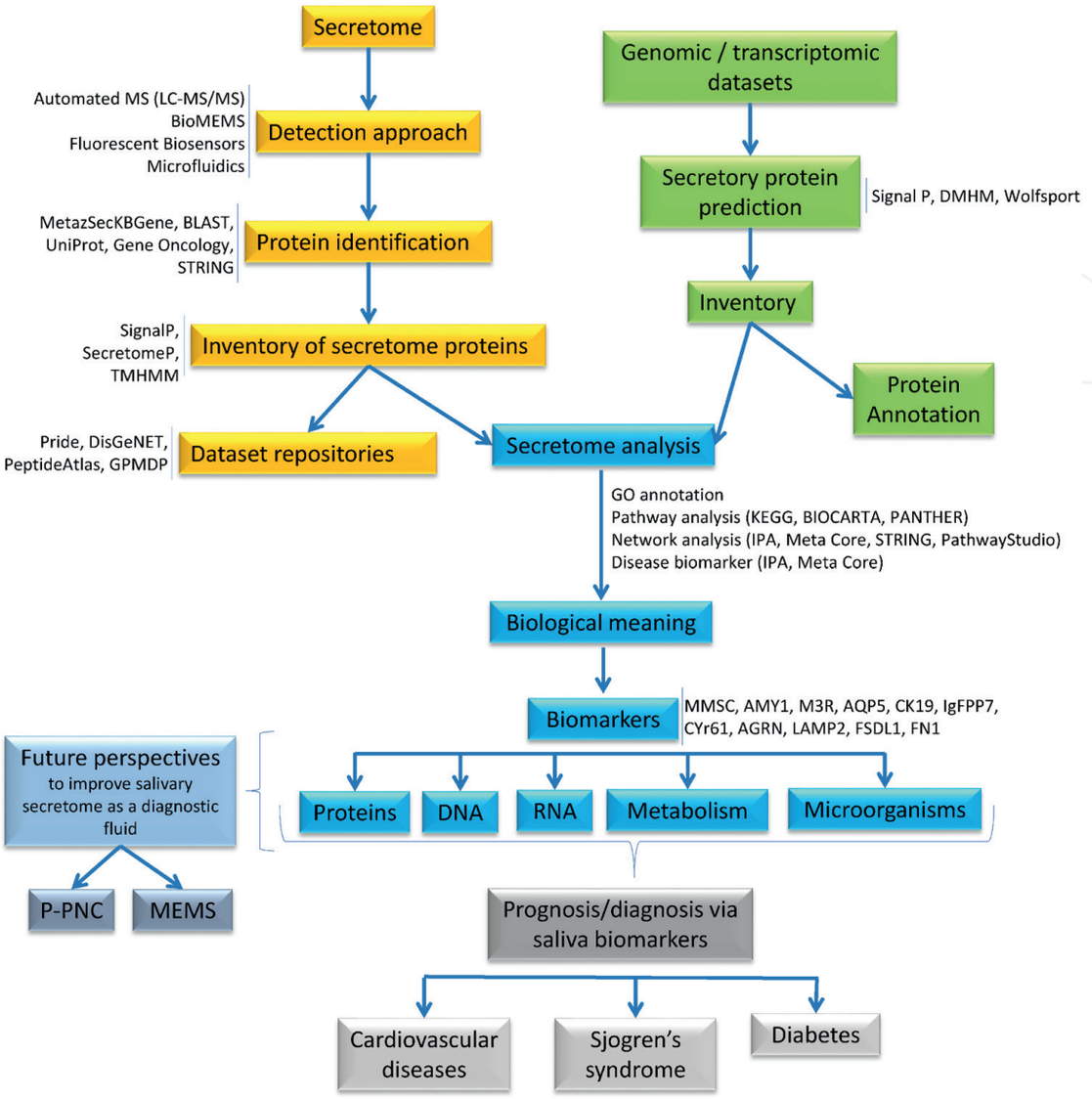

Figure 1.
Workflow of secretome analysis for the comprehensive characterization of molecules secreted by salivary glands, arriving at the final point (biomarkers) that can be used for prognosis/diagnosis in several diseases.

8. Conclusion

Although there are some limitations, salivary proteomics is a promising diagnostic and therapeutic tool for several critical diseases. Salivary gland secretome represents a valuable new tool to measure many local soluble mediators, provide future insight into immunopathology, and potentially aid in diagnosis.

Routine laboratory tests include hematology, clinical chemistry, and immunochemistry using high performance equipment. Diagnosis based on salivary secretome may provide an efficient, rapid and simple automated method for transformation. The next decade will bring improvements in accuracy, performance, and bed monitoring, but not hospital systems.

Improving basic healthcare systems with personalized medications, biosensors, lab-on-a-chip systems, personal genetics, and smartphone tracking parameters. The impact of saliva testing on healthcare systems is enormous, aggressive and convenient.

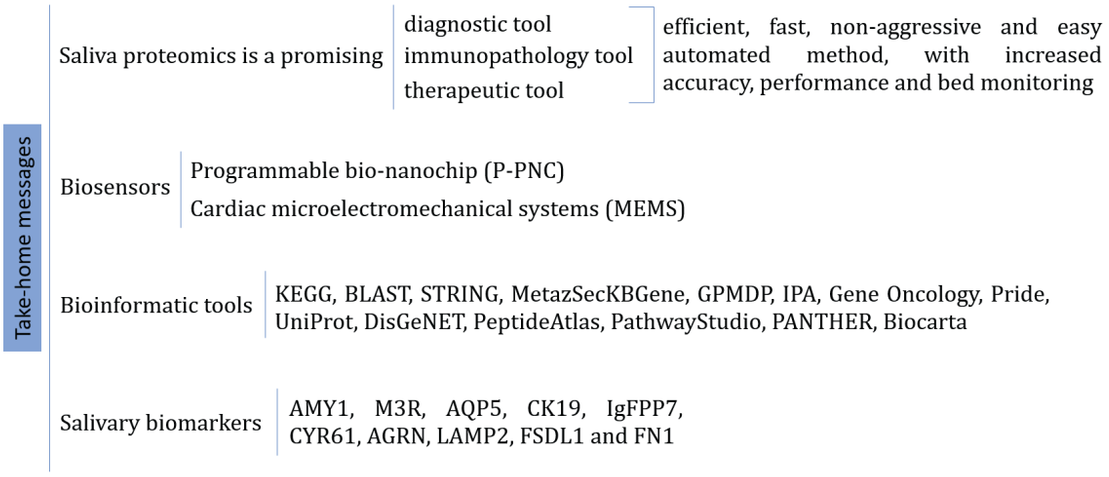

Figure 2.
Take-home messages that summarize the main ideas/concepts of this paper.

Reportedly, the ability to use saliva for a liquid biopsy is an important diagnostic tool for medical conditions and dental diseases. The simple model has information related to non-aggression and physical health, making it an attractive choice.

Some of the salivary secretome markers mentioned in this review are general markers, not specific to particular diseases. A more specific set of markers is needed to make salivary secretome an acceptable diagnostic fluid. The recent introduction of the programmable bio-nanochip system (P-PNC) has driven the revolution in saliva detection technology for the detection of cardiovascular disease (CVD).

Other biosensing systems, such as cardiac microelectromechanical systems (MEMS), can also be used to detect certain diseases. With the help of the latest labs in chip systems, they will improve hospital practice and human health [57]. Future development of this diagnostic tool will lead to further improvements in certain devices that will change the method of screening for critical diseases such as CVD.

A take-home message that summarizes, as shown in **Figure 2**, the main issues that have been addressed so far in salivary proteomics as a diagnostic and therapeutic tool. It also includes the means of detection and prediction of salivary proteomics (biosensors and bioinformatics tools). Although some have been used for a long time, most are novel tools and techniques that have been shown to provide great data to support proteomics studies. In addition, **Figure 2** provides a short list of the most promising and relevant salivary biomarkers discussed to date.

Acknowledgements

The authors thank the Portuguese Foundation for Science and Technology (FCT), European Union, QREN, FEDER and COMPETE for funding UnIC - Unidade de Investigação Cardiovascular (UIDB/00051/2020 and UIDP/00051/2020), iBiMED (UIDB/04501/2020, POCI-01-0145-FEDER-007628) and FCT LAQV/REQUIMTE (UIDB/50006/2020) research units. R.V. is supported by individual fellowship grants (IF/00286/2015). This work is funded by national funds (OE), through FCT Fundação para a Ciência e a Tecnologia , I.P., in the scope of the framework contract foreseen in the numbers 4, 5 and 6 of the article 23, of the Decree-Law 57/2016, of August 29, changed by Law 57/2017, of July 19.

DOI: http://dx.doi.org/10.5772/intechopen.98278

Conflicts of interest

The authors declare no conflicts of interest.

Author details

Luís Perpétuo[1], Rita Ferreira[2], Sofia Guedes[2], Francisco Amado[2] and Rui Vitorino[1,2,3]*

1 iBiMED, Department of Medical Sciences, University of Aveiro, Aveiro

2 LAQV/REQUIMTE, Department of Chemistry, University of Aveiro, Aveiro

3 Departamento de Cirurgia e Fisiologia, Faculdade de Medicina da Universidade do Porto, UnIC, Porto

*Address all correspondence to: rvitorino@ua.pt

References

[1] Khurshid Z, Zohaib S, Najeeb S, Zafar MS, Slowey PD, Almas K. Human saliva collection devices for proteomics: an update. International journal of molecular sciences, 17(6), 846 (2016).

[2] Dawes C. Salivary flow patterns and the health of hard and soft oral tissues. The Journal of the American Dental Association, 139, 18S-24S (2008).

[3] Podzimek S, Vondrackova L, Duskova J, Janatova T, Broukal Z. Salivary markers for periodontal and general diseases. Disease markers, 2016 (2016).

[4] Khurshid Z, Zafar MS, Khan RS, Najeeb S, Slowey PD, Rehman IU. Role of salivary biomarkers in oral cancer detection. Advances in clinical chemistry, 86, 23-70 (2018).

[5] Abdul Rehman S, Khurshid Z, Hussain Niazi F et al. Role of salivary biomarkers in detection of cardiovascular diseases (CVD). Proteomes, 5(3), 21 (2017).

[6] Saitou M, Gaylord E, Xu E et al. Functional specialization of human salivary glands and origins of proteins intrinsic to human saliva. bioRxiv, 2020.2002.2012.945659 (2020).

[7] Feizi A, Gatto F, Uhlen M, Nielsen J. Human protein secretory pathway genes are expressed in a tissue-specific pattern to match processing demands of the secretome. npj Systems Biology and Applications, 3(1), 22 (2017).

[8] Chansaenroj A, Yodmuang S, Ferreira JN. Trends in salivary gland tissue engineering: from stem cells to secretome and organoid bioprinting. Tissue Engineering Part B: Reviews, (2020).

[9] Hassaneen M, Maron JL. Salivary diagnostics in pediatrics: applicability, translatability, and limitations. Frontiers in public health, 5, 83 (2017).

[10] Pappa E, Kousvelari E, Vastardis H. Saliva in the "Omics" era: A promising tool in paediatrics. Oral diseases, 25(1), 16-25 (2019).

[11] Khurshid Z, Zohaib S, Najeeb S, Zafar MS, Rehman R, Rehman IU. Advances of proteomic sciences in dentistry. International journal of molecular sciences, 17(5), 728 (2016).

[12] Rahim MAA, Rahim ZHA, Ahmad WAW, Hashim OH. Can saliva proteins be used to predict the onset of acute myocardial infarction among high-risk patients? International journal of medical sciences, 12(4), 329 (2015).

[13] Wang X, Kaczor-Urbanowicz KE, Wong DT. Salivary biomarkers in cancer detection. Medical Oncology, 34(1), 1-8 (2017).

[14] Roi A, Rusu LC, Roi CI, Luca RE, Boia S, Munteanu RI. A new approach for the diagnosis of systemic and oral diseases based on salivary biomolecules. Disease markers, 2019 (2019).

[15] Meinken J, Walker G, Cooper CR, Min XJ. MetazSecKB: the human and animal secretome and subcellular proteome knowledgebase. Database, 2015 (2015).

[16] Silberring J, Ciborowski P. Biomarker discovery and clinical proteomics. TrAC Trends in Analytical Chemistry, 29(2), 128-140 (2010).

[17] Wang A, Wang CP, Tu M, Wong DT. Oral biofluid biomarker research: current

status and emerging frontiers. Diagnostics, 6(4), 45 (2016).

[18] Hsu C-W, Yu J-S, Peng P-H et al. Secretome profiling of primary cells reveals that THBS2 is a salivary biomarker of oral cavity squamous cell carcinoma. Journal of proteome research, 13(11), 4796-4807 (2014).

[19] Wei F, Wong D. Point-of-care platforms for salivary diagnostics. The Chinese journal of dental research: the official journal of the Scientific Section of the Chinese Stomatological Association (CSA), 15(1), 7-15 (2012).

[20] Soper SA, Brown K, Ellington A et al. Point-of-care biosensor systems for cancer diagnostics/prognostics. Biosensors and Bioelectronics, 21(10), 1932-1942 (2006).

[21] McRae MP, Simmons GW, Wong J et al. Programmable bio-nano-chip system: a flexible point-of-care platform for bioscience and clinical measurements. Lab on a Chip, 15(20), 4020-4031 (2015).

[22] Ngounou Wetie AG, Sokolowska I, Woods AG et al. Automated mass spectrometry–based functional assay for the routine analysis of the secretome. Journal of laboratory automation, 18(1), 19-29 (2013).

[23] Hosseini S, Espinosa-Hernandez MA, Garcia-Ramirez R, Cerda-Kipper AS, Reveles-Huizar S, Acosta-Soto L. BioMEMS.).

[24] Espinosa-Hernandez MA, Reveles-Huizar S, Hosseini S. Bio-microelectromechanical Systems (BioMEMS) in Bio-sensing Applications-Colorimetric Detection Strategies. In: BioMEMS. (Springer, 2021) 21-67.

[25] Naher S, Orpen D, Brabazon D, Morshed MM. An overview of microfluidic mixing application. In: Advanced Materials Research. (Ed.^(Eds) (Trans Tech Publ, 2010) 931-939.

[26] Morris MC. Fluorescent biosensors—probing protein kinase function in cancer and drug discovery. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1834(7), 1387-1395 (2013).

[27] Mitchell AC, Alford SC, Hunter SA et al. Development of a protease biosensor based on a dimerization-dependent red fluorescent protein. ACS chemical biology, 13(1), 66-72 (2018).

[28] Tu M, Cheng J, Chen Y-L et al. Electric Field–Induced Release and Measurement (EFIRM): Characterization and Technical Validation of a Novel Liquid Biopsy Platform in Plasma and Saliva. The Journal of Molecular Diagnostics, 22(8), 1050-1062 (2020).

[29] Wei F, Strom CM, Cheng J et al. Electric field–induced release and measurement liquid biopsy for noninvasive early lung cancer assessment. The Journal of Molecular Diagnostics, 20(6), 738-742 (2018).

[30] Mani V, Beduk T, Khushaim W et al. Electrochemical sensors targeting salivary biomarkers: A comprehensive review. TrAC Trends in Analytical Chemistry, 116164 (2020).

[31] Goswami Y, Mishra R, Agrawal AP, Agrawal LA. Salivary biomarkers: A review of powerful diagnostic tool. IOSR J. Dental Med. Sci, 14(3), 80-87 (2015).

[32] Yoshizawa JM, Schafer CA, Schafer JJ, Farrell JJ, Paster BJ, Wong DT. Salivary biomarkers: toward future clinical and diagnostic utilities. Clinical microbiology reviews, 26(4), 781-791 (2013).

[33] Gug IT, Tertis M, Hosu O, Cristea C. Salivary biomarkers detection: Analytical and immunological methods overview. TrAC Trends in Analytical Chemistry, 113, 301-316 (2019).

[34] Castagnola M, Scarano E, Passali G et al. Salivary biomarkers and proteomics: future diagnostic and clinical utilities. Acta Otorhinolaryngologica Italica, 37(2), 94 (2017).

[35] Sivadasan P, Gupta MK, Sathe GJ et al. Data from human salivary proteome–A resource of potential biomarkers for oral cancer. Data in brief, 4, 374-378 (2015).

[36] Proctor GB. The physiology of salivary secretion. Periodontology 2000, 70(1), 11-25 (2016).

[37] Keremi B, Beck A, Karoly Fabian T et al. Stress and salivary glands. Current pharmaceutical design, 23(27), 4057-4065 (2017).

[38] Wang J, Liang Y, Wang Y et al. Computational prediction of human salivary proteins from blood circulation and application to diagnostic biomarker identification. PLoS One, 8(11), e80211 (2013).

[39] Caccia D, Dugo M, Callari M, Bongarzone I. Bioinformatics tools for secretome analysis. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1834(11), 2442-2453 (2013).

[40] Ashburner M, Ball CA, Blake JA et al. Gene ontology: tool for the unification of biology. Nature genetics, 25(1), 25-29 (2000).

[41] Khatri P, Drăghici S. Ontological analysis of gene expression data: current tools, limitations, and open problems. Bioinformatics, 21(18), 3587-3595 (2005).

[42] Szklarczyk D, Franceschini A, Kuhn M et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic acids research, 39(suppl_1), D561-D568 (2010).

[43] Cline MS, Smoot M, Cerami E et al. Integration of biological networks and gene expression data using Cytoscape. Nature protocols, 2(10), 2366 (2007).

[44] Ma Z-Q, Tabb DL, Burden J et al. Supporting tool suite for production proteomics. Bioinformatics, 27(22), 3214-3215 (2011).

[45] Martens L, Chambers M, Sturm M et al. mzML—a community standard for mass spectrometry data. Molecular & Cellular Proteomics, 10(1), R110. 000133 (2011).

[46] Orchard S, Hermjakob H. The HUPO proteomics standards initiative—easing communication and minimizing data loss in a changing world. Briefings in bioinformatics, 9(2), 166-173 (2008).

[47] Deutsch EW, Lam H, Aebersold R. PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. EMBO reports, 9(5), 429-434 (2008).

[48] Jones P, Martens L. Using the PRIDE proteomics identifications database for knowledge discovery and data analysis. In: Proteome Bioinformatics. (Springer, 2010) 297-307.

[49] Smith BE, Hill JA, Gjukich MA, Andrews PC. Tranche distributed repository and ProteomeCommons. org. In: Data Mining in Proteomics. (Springer, 2011) 123-145.

[50] Rehman AU, Olof Olsson P, Khan N, Khan K. Identification of Human Secretome and Membrane

Proteome-Based Cancer Biomarkers Utilizing Bioinformatics. The Journal of membrane biology, 253, 257-270 (2020).

[51] Shin J, Rhim J, Kwon Y et al. Comparative analysis of differentially secreted proteins in serum-free and serum-containing media by using BONCAT and pulsed SILAC. Scientific reports, 9(1), 1-12 (2019).

[52] Min XJ. Evaluation of computational methods for secreted protein prediction in different eukaryotes. JPB, 3(5) (2010).

[53] Uhlén M, Fagerberg L, Hallström BM et al. Tissue-based map of the human proteome. Science, 347(6220) (2015).

[54] Keerthikumar S. A catalogue of human secreted proteins and its implications. AIMS Biophys, 3(4), 563-570 (2016).

[55] Chen G, Chen J, Liu H et al. Comprehensive identification and characterization of human secretome based on integrative proteomic and transcriptomic data. Frontiers in cell and developmental biology, 7, 299 (2019).

[56] Mona M, Kobeissy F, Park Y-J et al. Secretome Analysis of Inductive Signals for BM-MSC Transdifferentiation into Salivary Gland Progenitors. International journal of molecular sciences, 21(23), 9055 (2020).

[57] Khan RS, Khurshid Z, Yahya Ibrahim Asiri F. Advancing point-of-care (PoC) testing using human saliva as liquid biopsy. Diagnostics, 7(3), 39 (2017).