Faculty of Medicine and Dentistry

Doctoral Program in Dentistry 3143 RD 99/2011

Clinical and Basic Odontostomatology Research

DOCTORAL THESIS

INFLAMMATORY RESPONSE ELEMENTS, GLYCAN, AND PROTEOME PROFILES AS SALIVARY BIOMARKERS FOR THE EARLY DIAGNOSIS OF OSCC

Presented by

VALENTINA RADOSLAVOVA DIKOVA

Directed by

DR. José Vicente Bagán Sebastián

DR. Manuel M. Sánchez del Pino

Valencia, November 2020

Acknowledgements

The present doctoral thesis, beyond the science, research, and many hours of work, brings something as important as the experience shared with the people whose collaboration and support have contributed to its successful completion. Thereby, I would like to express my sincere acknowledgments to all those persons who walked along this way with me.

My deep gratitude goes first to Prof. José Bagán, for giving me the opportunity to carry out this study, for the trust and inspiration, for the guidance, immense knowledge, and experience he shared with me, contributing to my professional and personal growth. To Dr. Manuel M. Sánchez del Pino for the great expertise, unwavering enthusiasm for science, insightful comments, and encouragement during the project.

To the patients, for the volunteer participation and contribution to this scientific research.

To all the colleagues, medical doctors, nurses, and staff of the Research Foundation and the Department of Stomatology and Maxillofacial Surgery of the General University Hospital of Valencia, and the Division of Oral Medicine at the Dentistry Faculty of the University of Valencia for the help, troubleshooting and professional advice within the working and studying process, as well as for the friendship and great time spent together, in and outside the hospital and the lab.

To my colleague Sara, for being the most trustful and cheerful mate for sharing time of hard office and lab work, travels, events, and everlasting memories. To all my friends at home and abroad, for supporting me from near and far.

To Alexandra, for being the best friend I have ever had, and especially for the long conversations, new points of view, advice, and positive energy, always when I need it.

Last but not the least, I would like to thank my family: my brother, mum, dad, and grannies for always listening and making the effort to understand my scientific world, being proud of me and my work. For passing together my ups and downs, for the endless trust, and being a driving force to pursue my aims and dreams.

“Nothing in life is to be feared, it is only to be understood.

Now is the time to understand more, so that we fear less”.

-Marie Skłodowska-Curie-

Abstract

Oral cancer, predominantly oral squamous cell carcinoma (OSCC), is the most common malignant neoplasia in the oral cavity characterized by poor prognosis and a low survival rate when diagnosed with advanced disease. It can be treated effectively if detected at the initial stages with a better prognosis. Therefore, early diagnosis has become an ultimate goal in the management of oral cancer after prevention, considered to have a positive impact on the survival and quality of life of these patients. The shortage of symptoms at the disease onset often remains unnoticed, ultimately leading to discovery at advanced phases, so improvement of the diagnostic manner is required. Reliable biomarkers for OSCC are yet unavailable in the routine clinical setting, emphasizing the emerging need for a practical and simple tool to be used for definitive diagnosis, as well as for screening programs. OSCC tumors appear through a series of molecular mutations leading to uncontrolled cellular growth from hyperplasia areas to dysplastic lesions, to carcinoma *in situ,* and are finally followed by invasive carcinoma. Besides, the development of many OSCC cases has been correlated to the cancerous transformation of oral potentially malignant disorders (OPMDs) such as oral leukoplakia. The latter exhibits heterogeneous subtypes with varying modification potentials, among which proliferative verrucous leukoplakia (PVL) stands out with a high risk of malignant conversion. Diagnosis of OSCC requires sensitive and specific indicative tools that can support untraceable and difficult to access sites of the oral mucosa for which the most recommended discovery medium are biological fluids, such as saliva. Sputum-derived biomarkers offer easy sampling with reduced risk for the patients, cost, and diagnosis time, encompassing a variety of detectable and measurable parameters that can discriminate health from disease. In an attempt to identify novel salivary biomarkers of OSCC, we herein combined multi-omics analytical strategies to profile different types of molecules with potential diagnostic utility. As inflammation has previously been linked to oral pathologies, research so far indicates the possibility of using salivary pro and anti-inflammatory proteins for the screening of oral disorders. Altered cytokine responsiveness has been tightly associated with the development of OSCC, as well as detected in patients with OPMDs. To reveal changes related to oral carcinogenesis, a multiplex immune bead-based assay was used to survey

the levels of 8 different cytokines in the saliva of patients with PVL, at early and advanced OSCC stages and their healthy counterparts. Results indicated altered cytokine expressions associated with the pathology groups, the predictive models for sensitivity and specificity of which showed high ROC AUC values. It is now well established that defective glycosylation accompanies many chronic and infectious conditions and is a common feature of tumor cells that may affect *N*-glycans on glycoproteins. To compile a list of candidate biomarkers, another type of molecule has been studied. Saliva-derived *N*-glycans of healthy volunteers, PVL, and OSCC patients were analysed by the means of liquid chromatography (LC) coupled to mass spectrometry (MS). Comparative *N*-glycome profiling revealed several differentially expressed fucosylated bi- and tri-antennary glycans among the studied groups, providing a reasonable platform to further investigate the utility of salivary glycosylation for diagnosis of OSCC. Lastly, to get an insight into the complex molecular alterations associated with cancer, LC-MS generated proteomic profiles revealed more than 600 quantified proteins in the saliva of healthy and pre/cancerous lesions. The comparative analysis resulted in a list of differentially expressed proteins, characterized in OSCC, indicating significantly altered mechanisms implicating the immune system, inhibition of enzymatic activities, and cell adhesion. Among the identified candidate markers, some had previously been described in saliva. In addition, several newly OSCC-associated proteins have been annotated. In summary, in this discovery phase of biomarker identification, we provide a molecular panel of potential indicators of malignant transformation and/or early OSCC diagnosis. Further validation is needed to verify its clinical utility.

CONTENTS

I. *INTRODUCTION*

1. *Cancer Concept*

The generic term cancer unites a large group of diseases that can affect any part of the body and differ in their histogenesis, morphogenesis, clinical evolution, and prognosis. Other names used are malignant tumors and neoplasms. A defining feature of cancer is the rapid creation of abnormal cells that grow uncontrolled and disorderly beyond their usual boundaries. In consequence, it can acquire the ability to spread from the tissue of origin, through the circulatory and lymphatic systems, to adjoining tissues or organs giving rise to the process referred to as metastasizing. Metastases are a major cause of death from cancer.

2. *Head and Neck Cancer*

Head and Neck Cancer (HNC) is a term encompassing a group of biologically similar malignancies that arise in the upper aerodigestive tract. This anatomical region includes the paranasal sinuses, nasal cavity, oral cavity, pharynx, larynx, and salivary glands. According to the location, about 40% of the cancers occur in the oral cavity, 15% in the pharynx, 25% in the larynx, and the remaining tumors appear in other sites (salivary glands and thyroid) (1). More than 90 % of HNC are Squamous Cell Carcinomas (SCC) with heterogeneous histological patterns, emerging from the epithelium of these regions and can grow into deeper tissue layers. Non-squamous cell types are uncommon. Minor salivary gland carcinomas represent less than 5% of oral cavity cancers. They frequently arise on the hard palate (60%), lips (25%), and buccal mucosa (15%)(2). Mucoepidermoid carcinoma is the most typical representative salivary gland tumor, followed by low-grade adenocarcinoma, and adenoid cystic carcinoma. Other malignant formations can originate from connective tissue (sarcomas), lymphoid tissue (lymphomas), melanocytes (melanoma), or metastasis from distant tumors (3,4). HNC constitutes a major public health concern around the world and is a notable cause of morbidity and mortality. They are associated with high soreness due to interference with vital life functions such as breathing, swallowing, hearing, speech, vision, taste, and smell.

3. *Oral Squamous Cell Carcinoma*

Oral cancer is the most prevalent subtype of HNC that involves the lip and the intraoral mucosa. The oral cavity is divided into several anatomical sites: lip, tongue, the floor of the mouth, buccal mucosa, upper and lower gingiva, retromolar trigone, and palate (Figure 1). In the oropharynx area, we find the tonsils, the initial part of the pharynx, and the base of the tongue. Despite their proximity, the distinct anatomical characteristics of these sites have to be taken into account when planning oncologic intervention and/or therapy.

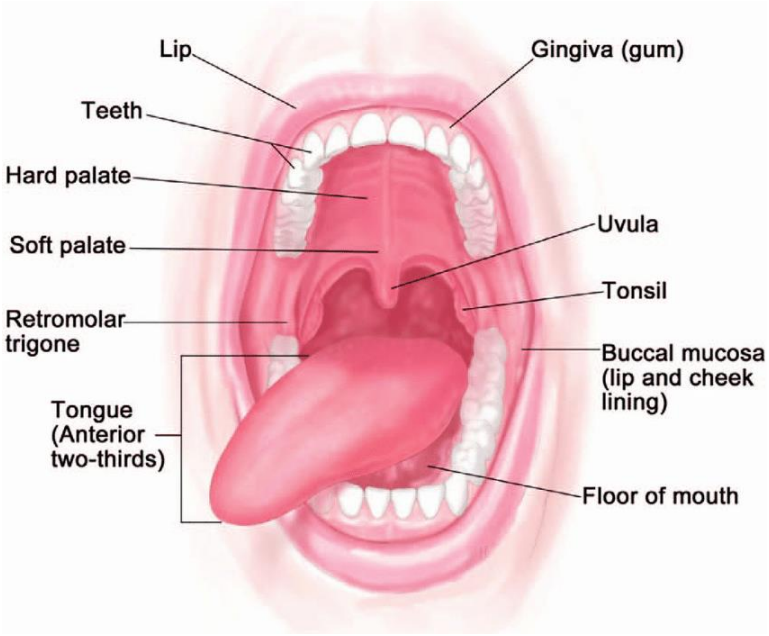


Figure 1. Oral cavity anatomical sites. Source: Terese Winslow LLC, Medical and Scientific Illustration (https://www.teresewinslow.com/)

Most of the oral epithelial neoplasia are from the squamous cell type (Figure 2) with the more representative type being the conventional Oral Squamous Cell Carcinoma (OSCC). Histologic classification is based on the subjective assessment of keratinization degree, cellular and nuclear pleomorphism, and mitotic activity, established by Pindborg and colleagues (5). According to it, well and moderately differentiated SCC tumors can be grouped as low grade and poorly differentiated as high grade. Verrucous Carcinoma (VC) is a non-erosive and non- metastasizing well-differentiated subtype with a good prognosis and indolent clinical course although the probability to progress to SCC. Basaloid SCC (BSCC) is a rare and aggressive high-grade kind occurring mainly in the larynx, hypo/oropharynx, epiglottis, and the base of the tongue. Despite the smaller number of

reported cases in the oral cavity, other variants comprising Spindle cell, Papillary, Adenosquamous, Acantholytic, Cuniculatum, and Lymphoepithelial subtypes belong to the categorization of SCC according to WHO and the International Agency for Research on Cancer (IARC) (6).

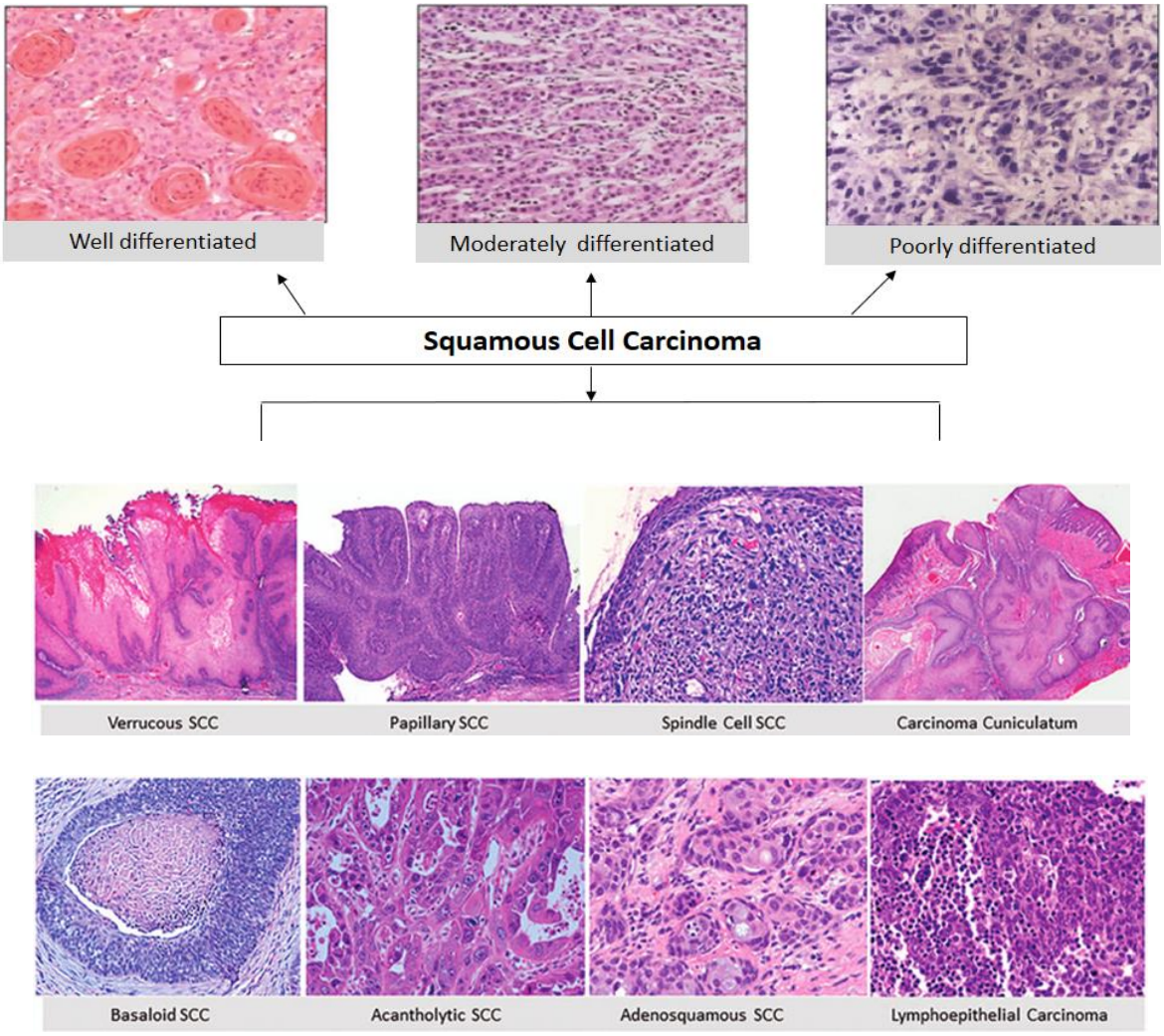


Figure 2. Morphologic features of SCC histologic subtypes according to WHO. Adapted after Pereira et al., 2007

3.1 *Clinical presentation*

A real danger of oral neoplasia is the asymptomatic onset when tumors can develop unseen. As the disease progresses, symptoms may vary from mild discomfort to severe pain, encompassing burning sensation, swelling, bleeding, ear pain, teeth mobility, dysphagia, complications using a prosthesis, and cervical lymphadenopathy (3). OSCC tumors can arise in apparently normal mucosa or be preceded by clinically visible premalignant lesions, most commonly found on the tongue and the floor of the mouth.

Other affected sites are the buccal mucosa, retromolar area, gingiva, soft and hard palate, and lip (7). Usually, the clinical presentation in the early stages is in the form of an erytholeukoplastic lesion comprising red and white areas with a slight roughness and relatively small size (Fig.3A). Typical for the advanced stages are lesion's increased size and profundity including ulceration, hardness, nodularity, and fixation to underlying tissues. In some cases may be presented as a lump with warty surfaces and poorly demarked borders (Fig.3B) (7).

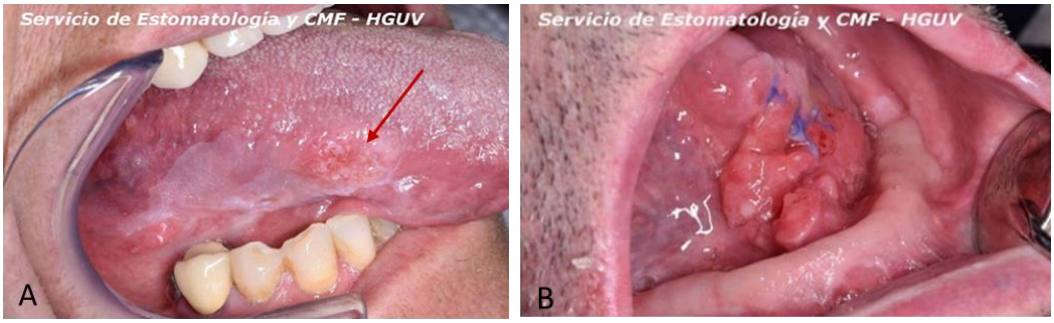


Figure 3. Early **(A)** and advanced **(B)** stage OSCC of the tongue.

3.2 *Epidemiology*

In the head and neck area, SCC is the most typical epithelial neoplasia that accounts for more than 90% of oral cancers and 2%-4% of all malignancies. In some regions of the Asian subcontinent the prevalence is higher, reaching 10% of all cancers in Pakistan, and around 45% in India (8). The oral cavity is the sixth most common anatomical location for cancer with over 350 000 new cases annually although the occurrence is highly variable, depending on the geographic location (9,10). The incidence of cancers of the lip and oral cavity are highly frequent in the Pacific Islands (Papua New Guinea), as well as Southern Asia followed by Australia/New Zealand and Europe (Fig.4). It is also the leading cause of cancer death among men in India and Sri Lanka (10). OSCC is usually a disease of elderly people among which males are affected twice as commonly as females, often diagnosed over the age of 50 years. Nevertheless, the incidence among young adults between 18 and 44 years is approximately 0.4 to 3.6% with a rising tendency (9,10). Over the past decades, the survival outcomes have not improved significantly and the prognosis remains relatively unfavorable, with 5-year overall and disease-free survival estimated to be 56% and 78%, respectively (6,7).

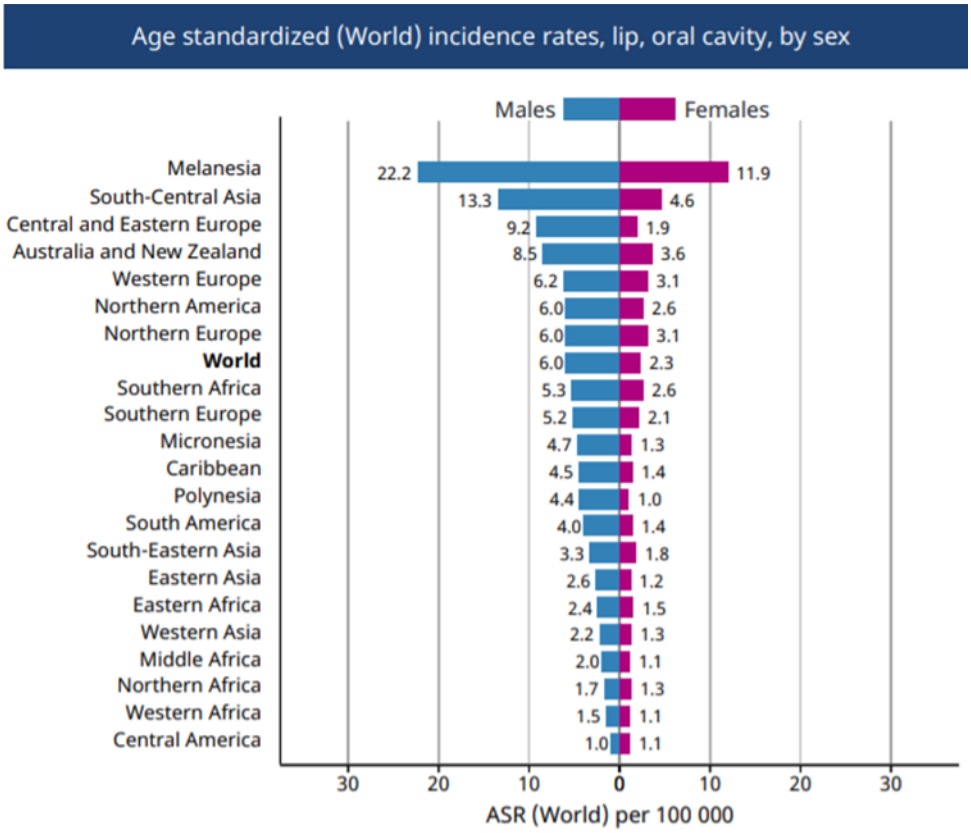


Figure 4. Bar chart representing Region-Specific, Age-Standardized (world) incidence rates calculated for males and females for cancers of the lip and oral cavity. Source: GLOBOCAN 2020.

3.3 *Etiology and Risk Factors*

Oral cancer develops through a multistage process whose mechanism of action remains not well understood, but the initial presence of a precursor cell subsequently evolving into cancer has been established. SCC tumors appear through a series of molecular mutations leading to the uncontrolled transformation of normal cells from hyperplastic areas to dysplastic lesions, to carcinoma in situ, and are finally followed by invasive carcinoma (11). These variations are taught to rise from the interaction between persons' genetic factors and environmental agents such as ultraviolet and ionizing radiation (physical carcinogenesis), asbestos, components of tobacco smoke, aflatoxin- a food contaminant and arsenic -a drinking water contaminant (chemical carcinogenesis) and infections from certain viruses, bacteria, or parasites (biological carcinogenesis). The microbiome denotes the collective genome of complex communities of bacteria, archaea, viruses, fungi, and protists, each with key roles to play in stabilizing microbial diversity (12,13). Exposure to tobacco, alcohol, and HPV infection has been shown to affect the oral microbiome resulting

in a shift towards heterogeneous bacteria (15-17). Moreover, the production of acetaldehyde by some bacterial species can induce mutagenesis and hyperproliferation of the epithelium (16). Within the diverse environment of the human mouth, distinct microbial populations can be found in both healthy and malignant sites, the clinical relevance of which lies in the statistical association between dysbiosis (often a result of poor oral health) and the prevalence of many cancer types (17). Alterations in the oral microbiome have been suggested to promote oncogenesis in the 7-15 % of oral cancer cases (18). The most common risk factor related to OSCC in the western world is considered to be the use of tobacco and alcohol, especially when acting synergistically (3). Together, they are associated with about 75% of all mouth and oesophageal cancers. Tobacco smokers carry a six-fold risk of oral cancer development compared to non-smokers. While it remains unclear whether alcohol intake alone is a causing agent of oral malignancies, collectively with tobacco are thought to be the greater reason for increasing incidence (19). Tobacco/areca nut/betel quid chewing habits spread between ethnic groups from the Taiwanese and Indian populations, as well as narcotics and cannabis consumption, are reckoned to be associated with the OSCC etiology (1,8). The role of some viral infections as potential disease causes is under investigation. In 2012, IARC disclosed that there was sufficient evidence to pertain Human Papilloma Virus (HPV) to oral carcinogenesis. In addition, HPV positive oral neoplasia differs in clinical response and overall survival rates from HPV negative tumors and cancers. According to the type and oncogenic potential, viruses have been classified as of low or high danger. In the mouth, 24 and 12 strains have been detected and linked to benign and malignant lesions, respectively. The majority of head and neck SCC infections are related to high-risk types 16, 18, 31, and 33 with HPV 16 being the prevalent one and HPV 33 accounting for about 10% in oropharyngeal (OPSCC) and OSCC cases. While the role of HPV in oropharynx and tonsil tumors is well defined and with higher frequency than in other head and neck parts, its infection of oral mucosal cells and predisposition to malignant transformation tends to be further studied (9,10]. Epstein–Barr Virus (EBV) is known as one of the most common human herpesvirus types also the cause of infective mononucleosis. Whereas its involvement in B cells malignant conversion is thoroughly described the contribution to OSCC pathogenesis remains uncertain, although it has been reported that LMP-1, the main oncoprotein of the EBV latent phase was expressed in 85% of OSCC cells (22). The action of Hepatitis C Virus (HVC), Torque Teno

viruses (TTVs), oral bacterial and fungal infections as co-carcinogens as well as genetic predisposition have also been discussed as increasing the risk factors for oral cancer development (23).

3.4 *Oral Potentially Malignant Disorders*

Most OSCCs are preceded by clinically visible but otherwise often asymptomatic lesions of the oral mucosa. According to the World Health Organization (WHO), lesions and conditions, that carry a risk for cancer are summarized as precancerous and considered as Oral Potentially Malignant Disorders (OPMD). A precancerous lesion refers to a benign, morphologically altered tissue having a potential for malignant transformation (11). After a workshop coordinated by the WHO Collaborating Centre for Oral Cancer in 2005, the consensus views of the working group were presented in a report reflecting a better understanding of multi-step carcinogenesis in the oral mucosa. It was recommended to relate “potentially malignant disorders” to pre-cancer as it conveys that not all disorders described under this term may transform into cancer rather that there is a family of morphological alterations amongst which some may have a high capacity for malignant conversion. OPMDs are taught to be indicators of risk of likely future neoplasia not only site-specific but elsewhere in (clinically normal-appearing) oral mucosa (24). Leukoplakia and Erythroplakia are the most common premalignant disorders, followed by Lichen Planus, Oral Submucous Fibrosis, Actinic cheilitis, Smokeless Tobacco Keratosis and palatal lesions in reverse smokers (25). The clinical term Leukoplakia should be utilized to recognize predominantly white plaques of questionable risk having excluded (other) known diseases that carry no increased danger for cancer (24). The lesions have a variable behavioral pattern and no specific histology. It may or may not show atrophy, hyperplasia, and/or epithelial dysplasia. Traditionally, leukoplakia is present in homogeneous and non-homogeneous variants additionally comprising subtypes, firstly, classified by Axel *et al*. and Pindborg et al. (5) and later adapted and refined by other working groups in international seminars. To date, the concept for Homogeneous Leukoplakia (HL) comprises a uniformly white, flat plaque/s (Fig.5A). Non-homogeneous are nodular (polypoid rounded red or white outgrowths), verrucous/proliferative verrucous leukoplakia (PVL) (wrinkled, rigid, white surface appearance) (Fig.5B), speckled - erythroplakia (red patches) (Fig.5C), and erythroleukoplakia (mixed red -and -white lesions) (Fig.5D). To distinguish between the

variations, clinicians are based on the surface colour and morphological characteristics, being also representative for the possible outcome. Lesions may transform from one type to another and may further progress to carcinoma *in situ* and invasive OSCC. Such changes have been described in follow-up studies by various authors (26–29) It is well appreciated that the risk of malignant conversion is greater within the non-homogeneous lesions rather than the homogeneous ones (24). In western countries, Oral Leukoplakia (OL) is described as the most common premalignant disorder with malignant transformation rates from 1-17% (30). With a greater tendency for cancerous conversion is considered to be its non-homogenous subtype PVL, reaching up to 70% (29,31). The etiopathogenesis and transformation mechanisms of OL remain unknown, though factors including tobacco and alcohol use, chronic inflammation, HPV, and EBV infection have been suggested as influential (32).

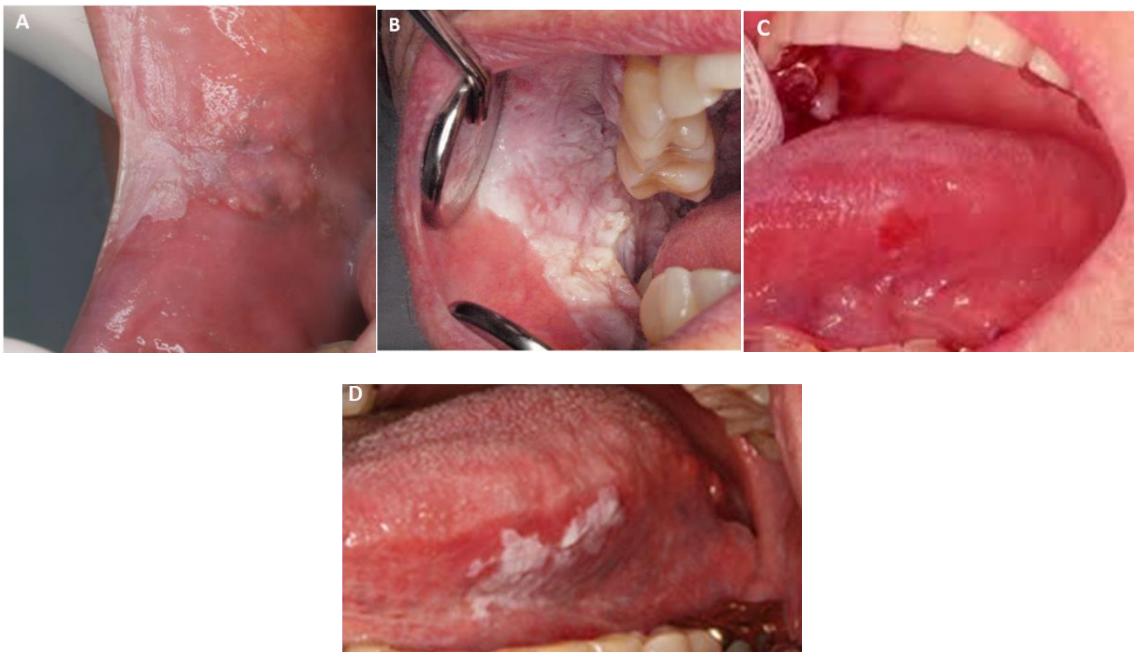


Figure 5. Homogeneous Leukoplakia (HL) **(A)** and Proliferative Verrucous Leukoplakia (PVL) **(B)** of the buccal mucosa. Erythroplakia **(C)** and Erythroleukoplakia **(D)** of the lateral border of the tongue. *Photographs obtained by the Service of Stomatology and Maxillofacial Surgery at the University General Hospital of Valencia.

Lichen planus (LP) is another pre-malignant mucocutaneous condition that can arise within the oral cavity. It is a chronic dermatologic disease without any specific etiology including drugs, C hepatitis, and nervousness. Its pathogenesis is a cell-mediated immune reaction to basal keratinocytes, more typical in middle-aged adults (33). The reticular form is

prevailing over the erosive, affecting the buccal mucosa bilaterally with interlacing white lines. Clinical manifestation of the erosive version is seen as an atrophic, erythematous lesion with central ulceration and a fine peripheral white area. Histopathologic features include different degrees of ortho- and parakeratosis besides the thickness of the spinous layer (34). Relatively low potential to SCC progression has been estimated (4%) (35). Typical for the Oral Submucous fibrosis (OSMF) lesions is a mucosal rigidity with different intensities caused by fibroblastic hyperplasia and modification of superficial connective tissue. Often involves young adults affecting the buccal mucosa, the retromolar area, and soft palate. Its histologic pattern reveals the fibroblastic transformation of the juxta epithelial connective tissue layer with a variable number of chronic inflammatory cells. Unlike other precancerous lesions, OSMF does not show regression, epithelial dysplasia has been noted in 10% to 15% of the biopsies with rates of malignant transformation to SCC between 2% and 8 % (34),(36). Actinic cheilitis (AC) is a chronic inflammation of the lip induced by compromised immunity and excessive exposure to solar or artificial UV radiation. Microscopically, the lesions present atrophic stratified squamous epithelium with remarkable keratin production and varying degree of dysplasia (34). AC is considered a premalignant disorder with a carcinogenic potential of about 3% (37). Smokeless tobacco keratosis (STK) is a condition that develops on the oral mucosa in response to direct contact with a sniff or chewing tobacco. Generally, it appears as a white-grey or white patch and becomes permanent unless the use of tobacco is avoided. Dysplasia, if exists, is mild in these lesions and is associated with a slightly increased risk of oral cancer (34).

3.5 *Diagnosis and Staging*

OSCC develops through a multistage process whose mechanism of action remains not well understood, but the initial presence of a precursor oral mucosal lesion subsequently evolving into a tumor outgrowth has been established. Many oral carcinomas are preceded by OPMD of which the non-homogeneous leukoplakia types are perhaps the most notable (24). There is no standard practice for patients presenting with clinically evident lesions to evaluate the risk of cancer. A common starting point is a conventional examination by a frontline clinician such as a general dentist. It is recommended a thorough visual check-up of the oral mucosa, mainly of the sites that are especially predisposed (Fig. 6) together with

palpation of the lymphoid tissue of the neck (cervical lymph nodes) for the presence of nodular formations (38).

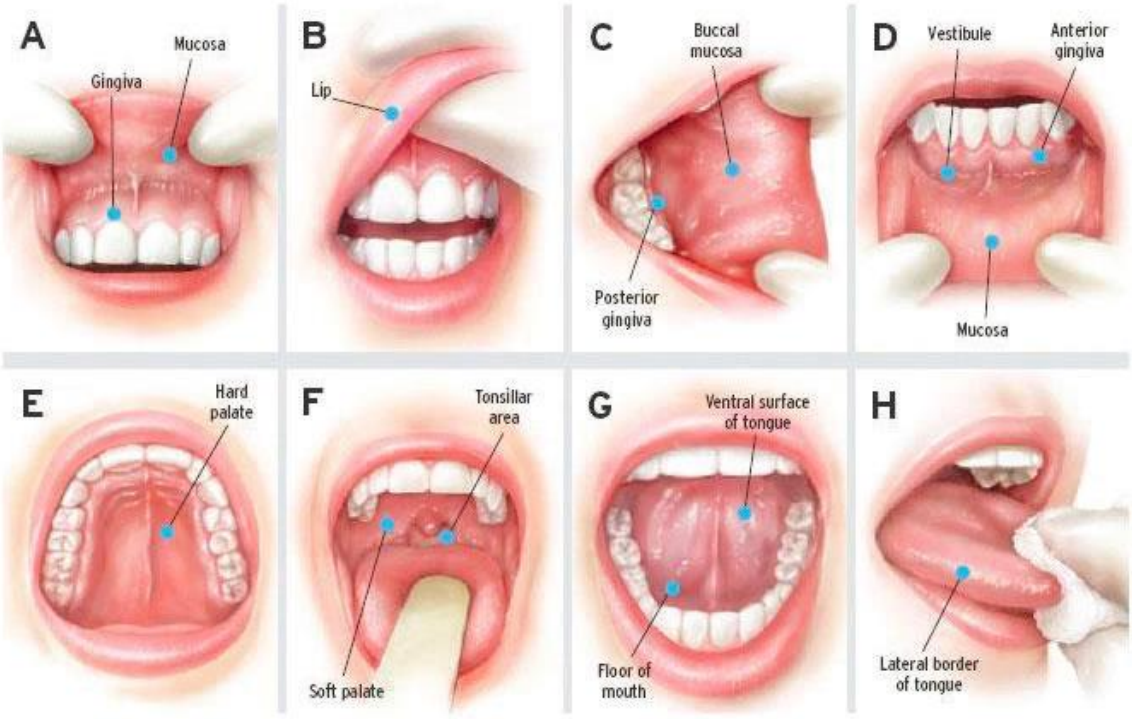


Figure 6. The 8-step oral cancer screening. Source: Cancer Society NZ

Upon discovering a suspicious lesion, persisting for more than 3 weeks, the clinician will make a subjective judgment based on the clinical presentation, experience, and resources available to decide the next step. Clear anticipation of malignancy is usual in advanced stages when fungating, ulcerative mass is obvious, frequently accompanied by pain, swelling, and radiographic destruction of bone. As early the disease spectrum is, as less evident the clinical features become and hence the precision of whether a lesion is, or has the potential to become malignant. Thus, the subjective assessment and the limitation of the clinical characteristics alone can easily result in the wrong diagnosis at the initial stages. To date, the gold standard for diagnosis remains the histopathologic assessment of an incisional tissue biopsy taken from a suspected area. An accurate histopathologic diagnosis depends on doing an appropriate biopsy and providing adequate clinical information (history of dysplasia or SCC, the patient's risk factors, lesion location, appearance, size, and duration), and on the correct interpretation of the biopsy results. Microscopically, active oral epithelial changes such as dysplasia, hyperplasia, hyperkeratosis, and acanthosis differ

from the normal pattern (Fig. 7). Oral Epithelial Dysplasia (OED) is a histopathologic term used to describe tissue changes observed in a chronic, progressive and premalignant disorder of the oral mucosa (39). Moreover, dysplastic alterations are consistently seen in the mucosa adjacent to the tumor in patients with invasive OSCC (Fig.7f). OED is considered the main histologic marker of possible evolution to malignancy and is thought to be predictive of an increased rate of SCC development (40). When transformations are limited to the basal or parabasal keratinocytes, the severity of dysplasia is described as mild (Fig.7c). The atypia affecting basal to the middle of the granular layer is referred to as moderate dysplasia (Fig.7d). Changes extending to the upper and middle layer of the epithelium are reported with the terms severe dysplasia and carcinoma *in situ* (atypia is complete from the base to the surface of the mucosa). In CIS the basal layer is intact, the dysplasia has not yet invaded into the subepithelium (Fig.7e) (34). It is generally characterized by the presence of a front formation (clear to obscure) against surrounding normal or dysplastic epithelial and dense lymphocytic infiltrations below the epithelial layer (41).

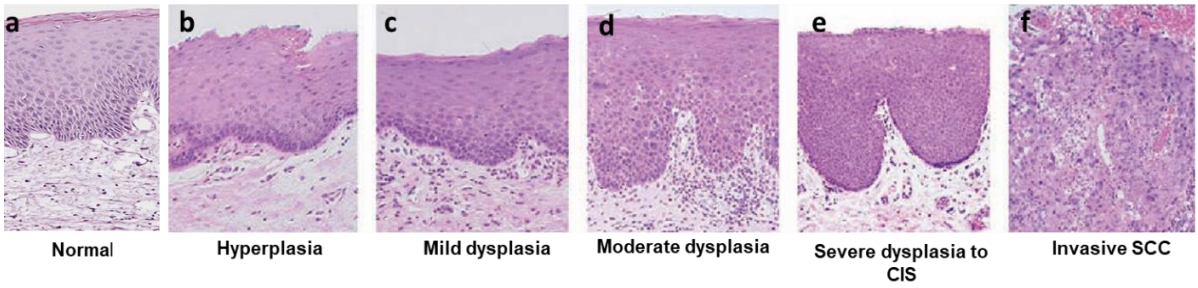


Figure 7. Normal epithelium (a) and histologic changes are believed to be associated with oral cancer progression from hyperplasia (b) to mild (c), high grade of dysplasia (d), to carcinoma in situ (e), and finally to invasive SCC (f). Modified after Cankaya H., 2015

Dysplastic changes are categorized into cellular and tissue types. Individual cell properties like shape, nucleus and nucleoli size, density, mitotic activity are important for the determination of specific alterations. Basal cell hyperplasia, hyperkeratosis, acanthosis, and atrophy are benign tissue appearances of major importance to diagnose and grade OED. Squamous cell hyperplasia results in thickened keratin layer (parakeratin and/or orthokeratin) of the oral mucosal epithelium (hyperkeratosis) and/or without thickened spinouts layer (acanthosis). The most common cause of hyperkeratosis is chronic irritation of regular friction on the soft tissues of the mouth. The presence of epithelial dysplasia in

oral lesions is predictive by a variable rate (6.6 to 36.4%) of transformation to invasive SCC (39),(42). However, the actual diagnosis of oral borderline malignancies ranging from epithelial dysplasia to early SCC stages has been considered too challenging. Investigators generally accept the fact that most oral cancer cases display considerable change before reaching such a state (34). Thus, dysplastic lesions with microscopic characteristic features could be predictive for possible malignant transformation. Among patients with a histopathological diagnosis of dysplasia, about 1/10 of the total may be in danger (43). For its evaluation, an objective assessment of the presence and severity of OED is needed and so a prediction for risk of cancer is markedly different for low-grade (mild or moderate) and high-grade dysplasia (severe dysplasia or CIS). Most low-grade cases do not progress to cancer; high-grade dysplasia, however, often progresses if left untreated (44). Early diagnosis and risk assessment of dysplastic alterations of OPMD requires a team effort from both clinicians and pathologists and is important for prevention and therapeutic procedures. Even though the histologic microscopic view of a biopsy is mostly reliable, some other techniques have been proposed as diagnostic tools for OED including brush biopsy, liquid-based cytology, toluidine blue, vizilite technique, and oral autofluorescence (34),(45). Nevertheless, the low specificity and sensitivity and the number of false-positive results of these approaches embarrass the definitive diagnosis. So that, they are considered as screening practices rather than diagnostic methods. Also, an immunohistologic technique utilizing mononuclear antibodies against specific markers of dysplastic cells can yield more accurate results in the determination of dysplasia type (34).Accurate staging is essential in the optimal selection of treatment modalities, extent, and prognostic outcomes. Staging of OSCC is performed using the (tumor-lymph node-metastasis) TNM classification system (46). The clinical staging (cTNM) of the oral cavity tumors is assigned for all cancer patients identified before the treatment and consists of primary characteristics, the neck, and assessment for distant metastases. The basic elements in the staging of the primary site are the tumor size and invasion of deep structures. cTNM is composed of diagnostic workup information based on clinical history and symptoms, physical examination, imaging and biopsy of the primary site, and other relevant examinations (46). Radiographic imaging is crucial for the assessment of regional lymph nodes as well as of the relation of the tumefaction to adjacent bone (3). Computer tomography (CT) scan is used for evaluation of bone and neck nodes, Magnetic resonance

imaging (MRI) provides complementary information about soft tissue extent and perineural invasion (3). A Positron emission tomography (PET) scan is useful for the estimation of distant metastases and for radiation therapy planning (3). Pathologic stage (pTNM) is given if the patient undergoes surgical resection and it is composed by information form the cTNM combined with operative findings and pathology review of the resected specimens (tumour histopathology and/or regional lymph nodes). Together with the tumour features such as degree of differentiation, infiltration and probability for recurrence form the definitive diagnosis allowing the selection of an appropriate therapeutic approach, postoperative practices and prognosis (8,47). Collectively, oral lesions with epithelial or mesenchymal origin may emerge in different sites of the mouth. Upon discovering of such a condition, the healthcare specialist has to make a subjective judgement based on the clinical presentation, their acumen and experience to decide on an adequate management. Predicting if and which epithelial precursor lesions will become dysplastic or even malignant and which will have an indolent clinical course is difficult. It is the clinician's responsibility to identify the correct timing and site of biopsy in such a "case-findings". Some factors associated with professional delay include small tumour size, treatment given prior to definitive diagnosis, increase in number of referrals prior to definitive diagnosis and referrals by general practitioners to specialist without clear description or suspicion of malignancy (48). More often, late diagnosis of oral pre/malignancies are due to patient delays. They do not realize the seriousness of the initial symptoms, which can go unnoticed due to the smaller area of the incipient stage and lack of discomfort/pain. Besides, certain sites in the oral cavity may not be visually accessible and wounds may be unseen. Clinicians should remain alert for lesions of suspicious etiology since early detection of malignancy and intervention can go a long way in the management of this physically and psychologically scarring condition (49).

3.6 *Treatment and Prognosis*

- **Therapeutic approaches of OPMDs**

Primary prevention is ideally the best method in the management of pre-malignancy. It is prudent to risk-stratify a patient and provide appropriate counseling and screening for higher-risk individuals considering their oral health status and personal habits. Secondary prevention is accomplished by early OPMD identification followed by conservative or non-

conservative management. Factors influencing the choice for therapy include patient risk factors for malignancy (age, sex, and habits) and lesion risk factors (classification, size, morphology, malignant transformation rate, and location). Conservative treatment of clinicopathologic diagnosed OPMDs consists of observation alone. Periodic screening can be appointed, based on the patient's risk stratification of developing OSCC and dental/medical compliance. Standardized follow-up protocols of patients who have existing OPMDs or previously excised lesions are yet missing. Determining whether a lesion will follow a stable and harmless development or will progress to invasive carcinoma is challenging with the routine histopathological diagnosis and has limited prognostic value.The time to malignant transformation is unpredictable and varies from months to years. New lesions can appear adjacent to existing or previously excised ones or in a different location. Patients with a history of premalignant/dysplastic lesions should be followed up over a long-term period. Eliminating high-risk and promoting protective behaviors (such as smoking and alcohol cessation and a healthy diet) is essential. It is extremely important, patients with OED be followed by a specialist, trained to assess and manage such lesions. Surveillance varies by clinician experience, patient and lesion risk factors, and clinical characteristics. In theory, medical intervention is available, given that there is a lack of approved therapies by any Food and Drug Agency (FDA) (43). Chemoprevention is the uptake of natural or synthetically manufactured compounds designed to stop malefic progression, acting directly on early neoplastic cells. Medication treatment modalities differ by the mechanism of action of the agent employed. Vitamin A, retinoids (beta carotene), and carotenoids might influence epithelial turnover. Steroids such as triamcinolone and clobetasol propionate are attributed due to their immunosuppressive actions. Epidermal Growth Factor (EGFR) inhibitors/antagonists prevent the activation of pathways contributing to the acquisition of malignant cellular phenotype. Anti-inflammatory drugs block cyclo-oxygenase activity and topical and intralesional agents such as bleomycin (a glycopeptide antibiotic) cause induction of DNA strand breaks (50–52). Retinoid therapies are mostly researched, despite they are limited only to the treatment of leukoplakia (53). Studies found that systemic retinoic acid and lycopene may be advantageous in terms of improvement in histological features (54). However, due to the toxicity rate of around 10%, relapse rate after stopping treatment of about 54% and the lack of long-term follow-up of patients, the effect of these

chemoprevention compounds has yet to be validated (52,53). Surgical removal is the invasive management of premalignant lesions. It can be performed with different techniques (scalpel, laser surgery, and vaporization, photodynamic therapy, cryosurgery) and may or may not be warranted. The rationale of excision is that removing clinically altered tissue the resection margins of which are wide and free of epithelial abnormalities could prevent the onset of oral cancer (55). Laser surgery is beneficial in terms of whole lesion removal, minimal damage to enclosing tissues, reduction of postoperative pain and edema, and minimal oral dysfunction (56). The use of CO_2 and Nd: YAG lasers has become a routine treatment of oral leukoplakia, mainly in two modalities: evaporation and excision. Evaporation is applicable at different stages, especially in patients with wide or multiple lesions, and causes limited postoperative discomfort. A drawback of the method is the unfeasibility of analysing the whole lesion histologically, in contrast to excision. The last, in turn, could lead to functional problems with large size lesions. A study from 2015 by Del Corso *et al* (55) concluded that surgical removal of OPMDs such as OL is recommended, particularly if clinical and histological features suggest any potential risk of evolution to OSCC. Yet, the better choice for managing non-homogeneous OL with any grade of dysplasia appears to be via CO_2 laser excision, in terms of recurrence. Nd: YAG laser evaporation is suggested for the removal of non-dysplastic, homogeneous lesions, and in wide anatomical sites where excision could cause patients disturbance (55). The implementation of cryosurgery was shown favorable over other procedures causing minimal/no pain, less swelling, no bleeding, simplicity, versatility, and low cost (57). Regardless of the operative modality, a possibility for future relapse and malignant development must not be excluded. Where resolution of a lesion is observed, recurrence and adverse effects are common. Medical science explores innovative approaches using stem cell treatment and tissue-engineered oral mucosa. In a relevant study, keratinocytes and fibroblasts isolated from dysplastic oral lesions were combined to provide a renewable source of epithelia. Convincing results pointed out that the mucosa has gained the ability to reproduce normal keratins with maturity and function (58). Comprehending the principles of tissue growth and application to produce a functional replacement for clinical utility is defined as tissue engineering (59). A common approach is the use of grafts and cultured epithelial cells via partial- or full-thickness tissue-engineered oral mucosa. The last represents a better simulation of the in vivo conditions, taking into account the native

anatomical structure of oral mucosa (60). Transplanted cells should adapt to the new environment and establish a correct function, though there is a risk of rejection of the engineered implant (61). Careful, lifelong follow-up examinations are an essential requirement for untreated lesions and after surgery. Considering the diverse endeavors of healing approaches, yet no one shows a benefit when compared to placebo. There is a lack of randomized controlled studies and the reviewed data are from series of retrospective studies and case reports only. All oral dysplasia must be followed up at least annually, even if the lesion has been completely resected (i.e., no clinically visible lesion remains), and regardless of whether the patient has eliminated habituated condition (tobacco, alcohol consumption, etc.). Increasing evidence shows that even when excision is clinically and histologically confirmed, molecular clones of altered cells may remain and later initiate further dysplasia or SCC. Despite the proposed preventive strategies for the management of OPMDs, to date, there is no reported proof of a reliably effective cure and better survival in case of malignant transformation (62).

- **Therapeutic approaches of OSCC**

Early diagnosis remains essential for effective treatment and favorable outcomes. Unfortunately, most OSCCs are diagnosed at advanced disease stages that require multimodal treatment including surgery, radiotherapy, and/or chemotherapy, and often involve interdisciplinary consultations with maxillofacial and reconstructive surgeons, dental professionals, otolaryngologists, histopathologist, behavioral therapists for lifestyle alterations, and others (4). Treatment modalities are still with high economic costs and vastly damaging alternatives. Surgical intervention is the therapeutic mainstay for respectable tumors, allowing accurate disease staging and providing information about the status of margins, tumor spread, and histopathologic features. In the early stages, it usually results in permanent care and can be used to select subsequent management based upon the assessment of risk versus benefit. If present, OSCC metastasis occurs in cervical lymph nodes in about 80% of the patients. Cervical lymphadenectomy (radical neck dissection) is appropriately carried out (4). Adjuvant postoperative treatment is assigned when there is a high risk of locoregional recurrence. This includes patients with locally advanced primary tumors (stage III and IV), positive surgical margins, bulky nodal disease, neck metastases, lymphovascular invasion, perineural invasion, and extracapsular spread (3). The application

of external beam radiation therapy is a traditional modality (63). Parceling radiation dose over time is performed via two approaches: hyper-fractionation and accelerated fractionation. In hyper-fractionation, small-dose fractions are delivered two to three times per day to take advantage of the more limited damage to normal tissue associated with smaller fractions. The accelerated fractionation aims to deliver treatment at a rate of weekly dose accumulation 20 to 50 % faster than that of standard fractionation. This acceleration substantially shortens the overall duration of treatment to reduce the potential for tumor re-population that may occur between fractions (64). The National Comprehensive Cancer Network recommends a definitive radiotherapy dose of at least 66 Gy (1 gray [Gy] is the absorption of 1J of energy per kilogram of tissue) to the primary cancer of the oral cavity, and also to the neck in the case of clinically evident cervical disease (33). Chemotherapy administration implicates cytotoxic agents like methotrexate, cisplatin, carboplatin, fluorouracil, paclitaxel, and docetaxel. Neoadjuvant or Induction Chemotherapy is applied before definitive therapy, whether that definitive therapy is surgery or drug treatment. Using chemo and radiotherapy simultaneously or in alternation is known as concurrent or Integrated Chemoradiation. It has been demonstrated that local and regional control and free survival of oral and oesophageal SCC patients with extracapsular spread and /or affected surgical margins improve significantly when irradiation is combined with cisplatin treatment (versus radiotherapy alone) (65–67). Regional or Intra-arterial Chemotherapy is a promising approach that delivers high drug levels to the tumor with less systemic toxicity. Robbins and colleagues (68) developed protocols for the administration of intra-arterial cisplatin in combination with radiation. Complete response in 75%, a partial response in 23%, and no progression in 2% of patients were reported. However, prior to the selection of concurrent chemoradiation, the risk of treatment-related complications should be assessed based on patients' characteristics like physiological age, comorbid conditions, etc. Increased incidence of hematologic, mucous-membrane, and gastrointestinal adverse effects related to chemotherapy should be taken into account, as well. Concomitant chemo-radiotherapy is best applied at centers where appropriate expertise and infrastructure are available (65). Invasive surgery is a considerably morbid procedure that can have debilitating consequences affecting patients' physiological functions like chewing, swallowing, speech impairment, or facial aesthetics. Adverse side effects can also occur as a result of radiotherapy since there is a

liability to xerostomia, mucositis, and osteonecrosis (8). Restoration of integrity and function after ablative surgery is the final goal of treatment, achieved by choosing an appropriate reconstruction procedure. Defects after resection of early-stage tumors can usually be reconstructed with primary closure or the use of skin graft while recovery after the removal of advanced neoplasia requires the participation of an expert reconstructive surgeon (4). Recently, targeted molecular therapy has been applied to oral cancer patients (69). Treatment modality with monoclonal antibodies (targeted immunotherapy) has limited or non-existing side effects on normal cells, unlike chemo and radiotherapy. Studies including immunotherapeutic approaches and inhibitors of EGFR, VEGFR, IGF-1R, PI3K/AKT/mTOR, and MET have been conducted and expected in a near future to direct the management of oral cancer patients towards more personalized treatment (70). Strategies have focused mainly on four molecules associated with OSCC proliferation and differentiation: EFGR, cyclooxygenase-2 (COX-2), peroxisome proliferator-activated receptor γ (PPAR γ), and progesterone receptor (69). The admission of Cetuximab (Erbitux) in the clinical practice was a great advance in the field of targeted drugs. This chimeric monoclonal antibody targeting EGFR with high specificity and affinity has been the only FDA-approved drug ratified for both HPV- positive and negative subtypes. However, not all patients treated with Cetuximab respond well to therapy due to primary or acquired resistance, limiting significantly its efficacy. Several new immunotherapeutic agents have been introduced into clinical trials to target EGFR or its signaling partners (Panitumumab, Zalutumumab, Nimotuzumub, Afatinib, Erlotinib, Sym004). So far, they have yielded only modest improvements in progression-free survival and none have been affirmed for treatment in OSCC (9). Other recently approved molecular targeted intermediaries are immune checkpoint inhibitors being the co-signaling of the programmed cell death protein 1 (PD-1) and its ligand (PD-L1). Immunotherapy with other Pembrolizumab and Nivolumab, either as monotherapy or in combination with chemotherapy offers a new standard of care for patients with recurrent or metastatic PD-L1 positive tumors. The primary analysis showed safety and stabilized patient-reported quality of life. Besides, long-term follow-up treatment with Nivolumab demonstrated significantly improved overall survival (71),(72).

- **Prognosis**

The prognosis of OSCCdepends on several well-known critical factors associated with survival like disease phase at initial diagnosis, neck metastasis, invasiveness of cancer cells, and tumor thickness, etc. The clinical stage at presentation and the appearance of metastatic lymph nodes are important predictors of survival (Fig.8). It is generally accepted that prognosis is favorable in the early stages (I and II), especially when SCC are well-differentiated and not metastasized. A good outcome, with cure rates of 78% (stage I) and 68% (stage II) is expected for these cases (73). Unfortunately, a significant number of patients present with advanced disease stages (III or IV) at the time of the diagnosis when survival at 5 years is less than 50% and a cure of about 30% (3,73). Powerful histopathologic predictors of outcome include depth of invasion of the primary tumor, positive margins of surgical resection, perineural invasion, and major extracapsular nodal extension (3). Untreated individuals with metastatic disease show survival of about 4 months (74). Despite advances in treatment modalities and targeted therapy, the overall outlook for OSCC patients remains poor, due to invasion, metastasis, and high risk of local-regional recurrence. Approximately one-third of the subjects treated for oral cancer relapse and development of subsequent new primary tumors is the most common pattern of failure. OSCC presents with a recurrence rate of about 32% and 40-50% are with advanced disease relapse (75). This discrepancy can be explained both by the lack of effective chemoprevention and radiotherapy resistance (3). The possibility of second SCC primary ranges from 4–7% yearly (76). Close follow-up comprising thorough clinical examinations with high suspicion together with control of lifestyle-related risk factors remain the most important tactics for prevention and survival.

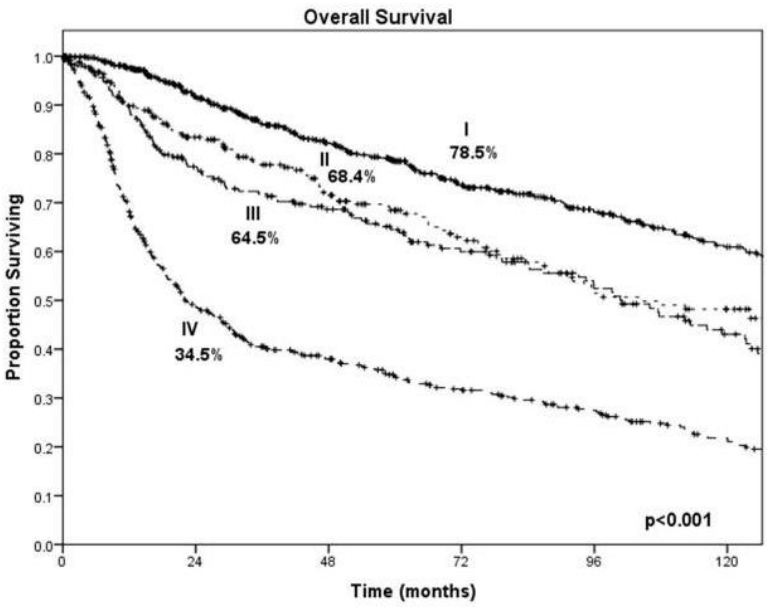


Figure 8. OSCC clinical stage at presentation as a prognostic factor. Source: Montero *et al.* 2015; Courtesy of Memorial Sloan-Kettering database, New York, NY.

4. *Molecular pathogenesis of OSCC*

Oral cancer is a heterogeneous disease that develops through a complex, stepwise process involving genetic changes, growth regulation, apoptosis, immortalization, angiogenesis, invasion, and metastasis. This process is accompanied by multiple molecular events resulting from the combined influence of an individual's genetic predisposition and exposure to environmental carcinogens. Squamous tumourigenesis is thought to result from the successive accumulation of molecular genetic alterations in the squamous epithelium lining of the upper aerodigestive tract. Accumulation of such alterations in epithelial cells precedes phenotypic changes, in many cases, associated with pre-invasive dysplastic lesions. The progression of the late stage of severe dysplasia to invasive carcinoma is comprised of both cellular and structural modification as a result of dysregulation of key pathways triggered by the interaction of epithelium and the host stromal elements (77).

4.1 *Genetic alterations during OSCC development*

A genetic progression model has been developed by Califano, Sidransky, and colleagues based on their studies of gene alterations in squamous cell carcinomas of the head and neck (SCCHN)(78,79). They found that the most common genetic aberration is the loss of

chromosomal region 9p21, which occurs in 70–80% of dysplastic lesions of the oral mucosa, suggesting that this loss is an early event in oral carcinogenesis. Another typical early genetic change is the loss of the chromosome 3p region (80,81). Loss of heterozygosity (LOH) of chromosome region 17p and mutation of the p53 gene occur in the later stage of progression from dysplasia to invasive squamous carcinoma. Alterations of p53, including mutation or deletion, are associated with increased genomic instability in oral dysplasia and may accelerate the rate of genetic alterations in oral carcinogenesis (81). Amplification of 11q13 and overexpression of cyclin D1 have been described in 40% of cases of oral squamous dysplasia (82). In general, loss of chromosomal material at 9p, 3p, and 17p is observed in relatively high proportions of dysplastic lesions, indicating that those events are early markers of oral carcinogenesis, whereas losses at 13q and 8p are observed more frequently in carcinomas than in dysplasia and are associated with later stages of carcinogenesis (79).

4.2 *Field cancerization*

In 1953 Slaughter coined the concept of "field cancerization", based on extensive histologic examination of dysplastic epithelium adjacent to invasive oral cancers (83). According to this theory, since the oral epithelium may be exposed to various carcinogenic insults, the entire area is therefore at increased risk for malignant transformation from the accumulation of genetic mutations of onco- and tumor suppressor genes. It has been developed from the finding of dysplastic epithelium adjacent to invasive oral cancers, which accounts for the high incidence of second primary tumors in OSCC treated patients. Many of these second primary tumors are associated with a lower rate of survival than the original tumor. Califano et al. (79) have also described the genetic progression model of field cancerization in head and neck cancer. In it, the transformation of normal mucosa is initiated by hits to the 9p chromosomal region. The resulting benign hyperplasia, further transforms into dysplasia due to following mutations in the 3p and 17p locations, while ultimate modifications in 17p are suggested to give rise to the first patch of mutated neoplastic cells. The patch further expands and converts to cancer along with mutations in 11q,13q, and 14q chromosomes (79). An important clinical implication of the genetic progression model is that fields often remain after surgery of the primary tumor and may lead to new cancers, as second primary or local recurrences (89). With increasing evidence

indicating cancer stem cells (CSCs) being a subgroup of cells in charge of tumor initiation and recurrence, it has been questioned whether the genetic changes characteristic to field cancerization are initiated within the resident stem cells of the mucosa. The origin of these cells is explained by various processes. One describes that tissue-specific stem cell undergoes several genetic and epigenetic alterations to give rise to a CSC (84). Another proposal is that CSCs originate from stem cells that acquire a premalignant phenotype during the developmental stage itself (84),(85). A third model, states that the CSCs originate from mature tumor cells that experience de-differentiation into a stem cell through alterations in signaling pathways and regulatory mechanisms (86). Besides, another probable origin of CSCs is the de-differentiation of mature oral epithelial cells (84,85). In oral mucosa where differentiated epithelial cells are high renewable (14 days) (87), it is presumable that normal stem cells (NSCs) as long-term residents of the epithelium have a higher possibility to accumulate genetic hits, requisite for malignant transformation. To reconcile these findings, Simple et al. have proposed a model of field cancerization orchestrated by the CSCs (88). As per it, independently of the cancerization process, the multistep molecular progression model of field cancerization may be CSC-driven (Fig. 9). Furthermore, second or multiple cancers distant from the dysplastic fields can be clonally related and derived from the expansion of a common pre-neoplastic progenitor. According to this proposal, a stem cell located in the basal epithelial layer acquires a genetic change - TP53 gene mutation (17p) and subsequently gives rise to daughter cells (a clonal unit) whereby all share the DNA alteration leading to the transformation of the normal stem cells (NSCs) into a patch of transit-amplifying cells (TACs) (Fig.9i). This transformation is accompanied by additional genetic mutations in chromosomes 3p and 9p. The patch forms an expanding field to the surrounding oral mucosa. The resultant mucosal field pushes the normal epithelium aside and may expand to a size of several centimeters (Fig.9ii). In some instances, the cells in this stage present dysplasia and may appear as leukoplakia or erythroplakia but often remain clinically undetectable. A subsequent 13q gene mutation where the location of the Rb gene is believed to release the CSCs from their lethargic stage hereby leading to proliferation, self-renewal, and primary tumor formation. Chromosome 11q13 gene amplification has been implicated in the progression of the field to cancer (Fig.9iii) which happens either by the mono- or by polyclonal nature of cancerization. In the monoclonal mode, the CSCs of the field can migrate laterally (intra-epithelial) to spread or

get implanted at a new site (inoculation through saliva) and form a genetically similar tumor at a later stage. Conversely, in the polyclonal field cancerization, NSCs at distinct sites in the mucosa, undergo a successive transformation into CSCs through independent, carcinogen-mediated molecular changes. This CSCs proliferate leading to the development of clones at different locations. Further genetic hits give rise to a subsequent aberration in the sub-clones within the field where the clone that ultimately gets the final mutation at 13q develops into carcinoma. The likelihood of the latest transforming event happening in a patient depends on the extent of exposure to carcinogens and the number of stem cells transformed afterward (9,90).

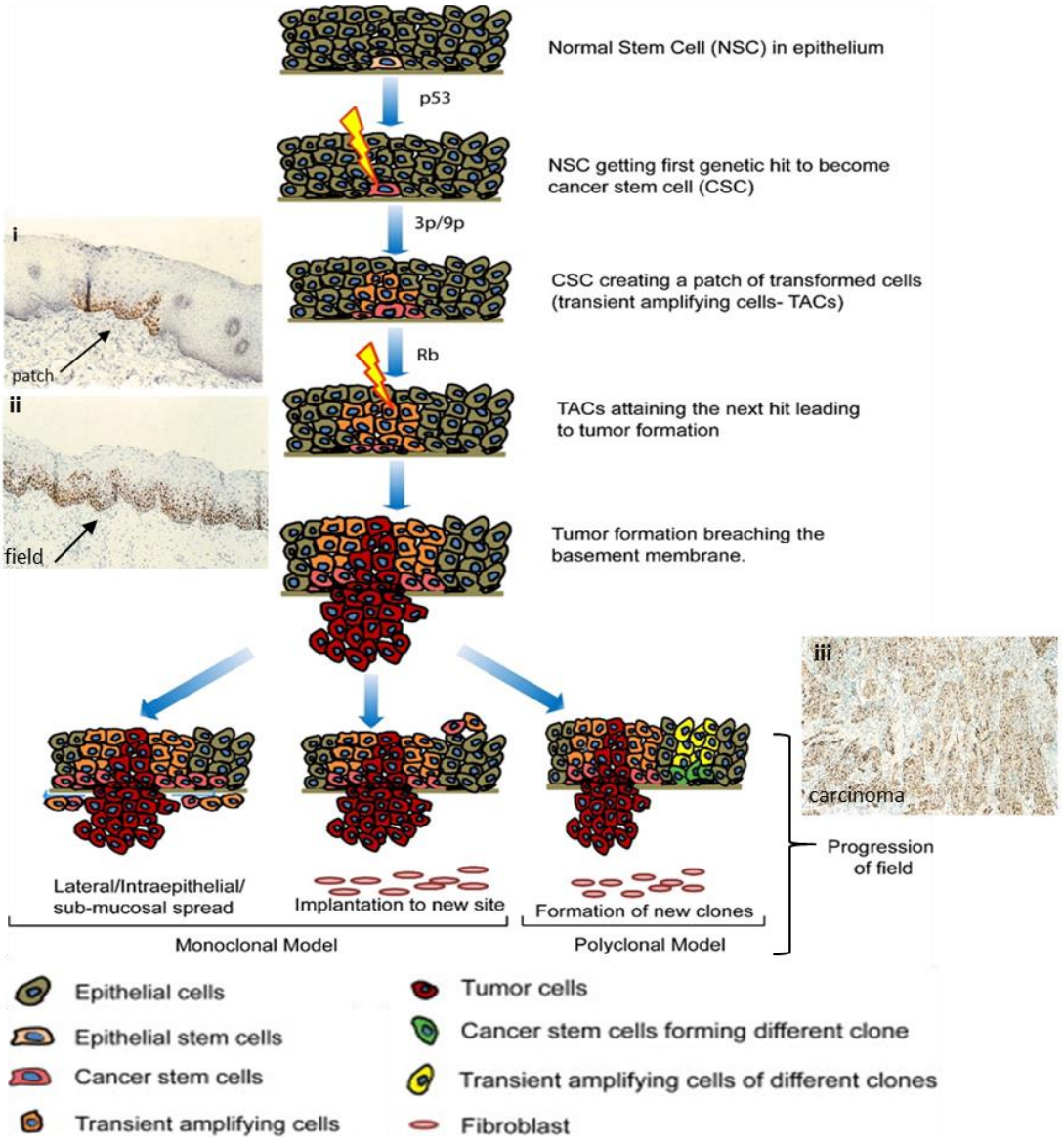


Figure 9. Proposed cancer stem cell-driven field cancerization model for cancer in the head and neck mucosa. As per it, the process of field cancerization is initiated by carcinogen assault leading to a TP53 mutation in a stem cell residing in the normal epithelium. The proliferation of this yet cancer-like stem cell results in a patch formation **(i)** accompanied by 3p and 9p chromosomal aberrations. The next step is the conversion from a patch to a field **(ii)**, an epithelial lesion consisting of cells with cancer-related genetic alterations, which grows at the expense of normal tissue. After a subsequent genetic hit to the 13q chromosomal region, one of the cells in the field will form the primary tumor. Other genetic changes take place for the progression from field to carcinoma **(iii)** which occurs either by the monoclonal or polyclonal mode of cancerization. Adapted after Braakhuis et al. (91) and Simple et al. (88).

Further research focused on the transcriptional changes that occur in the progression from normal-appearing mucosa to dysplastic tissue and invasive SCC (90). Microarray analysis of RNA isolated from a continuum of such specimens revealed genes with differential expression patterns. This transcriptional model demonstrates that the majority of alterations occurs before cancer development and that the difference between premalignant and malignant state is relatively little, as a comparison between a normal and premalignant state when a greater proportion of changes occurs. This data is consistent with previous genetic progression models derived from DNA-based mutations (79). Identified genes that were significantly upregulated in the progression process include integrin α 6 and GAPDH. Like other cancers, OSCC is also believed to arise through a complex process involving activation of oncogenes as well as the inactivation of tumor suppressor genes (77).

4.3 *Proto-oncogenes and oncogenes*

Proto-oncogenes are normal genes coding for proteins that help to regulate cell growth and differentiation. Proto-oncogenes are often involved in signal transduction and execution of mitogenic signals, usually through their protein products. Upon acquiring an abnormal activation of only one of the two gene copies, a proto-oncogene becomes a tumor-inducing agent, an oncogene that can promote uncontrolled cell growth, and proliferation, leading to tumorigenesis. Some of the common mechanisms of oncogene activation include mutation, chromosomal translocation, gene amplification, and retroviral insertion. Oncogenes can be classified into five groups based on the functional and biochemical properties of the protein products of their normal counterparts (proto-

oncogenes). These groups are: (1) growth factors or growth factor receptors (hst-1, int-2, EGFR/erbB, c-erbB-2/Her-2, sis), (2) intracellular signal transducers (ras, raf, stat-3), (3) transcription factors (myc, fos, jun, c-myc), (4) cell-cycle regulators (cyclin D1), and (5) those involved in the inhibition of apoptosis (bcl-2, bax,mdm2) (78). Oncogenes drive abnormal cell proliferation as a consequence of genetic alterations that either increase gene expression or lead to uncontrolled activity of the oncogene- encoded proteins.

4.4 *Acquisition of self-sufficient growth-stimulatory signaling*

Normal cells require exogenous growth signals to stimulate proliferation. Growth stimuli include soluble and membrane-bound growth factors, interactions with the extracellular matrix, and cytokines (92). Typically, these growth signals are transduced from cell-surface receptors that subsequently activate multiple intracellular signaling pathways, resulting in cell proliferation. During oral carcinogenesis, growth signaling can become dysregulated through an elevated level of growth factor receptors and/or their ligands, to promote autocrine stimulation without exogenous factors (78). Several intracellular growth signal-transducing proteins that are downstream mediators of growth factor signaling are frequently altered in cancer. These are described in Table 1 in the context of OSCC development and progression.

Table 1. Summary of oncogene-encoded regulators and their involvement in the molecular pathogenesis OSCC.

Oncogene-encoded regulators	Description	Signaling	Association with OSCC
Epidermal Growth Factor Receptor (EGFR)	EGFR (HER1 or ErbB1) is a transmembrane receptor, a member of HER/ERbB family of receptor tyrosine kinases (RTKs) which also includes ErbB2, ErbB3, and ErbB4. High-affinity EGFR ligands are EGF, TGFα, HBEGF, and BTC, whereas AREG, EREG, and EPGN constitute low-affinity ligands (93). The most common alteration	EGFR forms a dimer with another EGFR molecule and these receptors autophosphorylate, leading to a cascade of intracellular signaling events, including activation of Ras/Raf/MAPK, PI3K/Akt/mTOR, Jak/STAT, and PKC pathways. These pathways mediate multiple functions, including cell proliferation and survival, invasion, metastasis, and angiogenesis (94,95).	EGFR overexpression is reported in 80-90% of HNSCC with a mutation of EGFRvIII detected in about 42% of the reported cases (96). Higher EGFR expression was estimated in 70% of OSCC patients (97), increasing progressively from oral premalignant lesions to invasive OSCC (98,99). Besides, it has been demonstrated that EGFR amplification

	of EGFR is a truncation mutation, EGFR variant III (EGFRvIII) (77).		was associated with the advanced OSCC clinical stages, negatively impacting the disease progression (100).
C-MET and Hepatocyte Growth Factor/Scatter Factor(HGF/SF)	C-MET encodes c-Met factor which is a receptor tyrosine. Its only known ligand is (HGF/SF). c-Met is normally expressed by cells of epithelial origin, while expression of HGF/SF is restricted to cells of mesenchymal origin (77).	Stimulation of c-MET via its ligand in the MET pathway leads to numerous biochemical and biological effects, including increased cell migration, angiogenesis, proliferation, invasion, and metastasis (101).	c-Met mutations in HNSCC are rarely reported (2-13%), while the gain in C-MET gene copy number and HGF overexpression are common and often correlated with worse prognosis and lower overall survival (102). (103). HGF overexpression due to C-MET activation has been detected in primary OSCC tumors, supporting that the MET oncogene is involved in OSCC progression towards an invasive-metastatic behavior (104).
RAS oncogene	RAS is a proto-oncogene found to be involved in cell growth regulation and the transduction of mitogenic cell signaling from the cell surface to the nucleus. The three RAS genes (H-Ras, K-Ras, and N-Ras) are the most common oncogenes in human cancer (105).	All RAS encoded proteins belong to a large family of small GTPases (G-proteins), involved in cellular signal transduction. Two of the main cellular pathways in which RAS protein operates are the MAPK and PI3K pathways. MAPK comprises four sub-pathways of which the Erk1/2 pathway has received the most attention in oral cancer (106).	In oral cancer, a high incidence of H-Ras mutation has been found, mainly in Asian populations, where it has been correlated with betel nut chewing (107). Nevertheless, H-Ras mutations are found in about 5% of the OSCC cases in the West (108), and other RAS genes are infrequently mutated in OSCC (109). Aberration in the MAPK pathway represents only 4% of cases (110).
Cyclin D1	Cyclin D1 (CCD1) is a proto-oncogene encoding a positive regulator of G1	The binding of CCD1 to CDK4 or CDK6 leads to the phosphorylation of pRb subsequently triggering the release of transcription factors to	Overexpression of the cyclin D1 gene has been reported in 25-70% of oral

	phase progression through the cell cycle that regulates the initiation of DNA synthesis. It's also intricately involved in the regulation of apoptosis (111).	allow transcription of genes required for the progression through the G1 phase of the cell cycle in the absence of extracellular mitogen stimulation. Consistent with this function, cyclin D1 overexpression results in a more rapid transition from G1 to phase S. (112). The last is a common event in cancer but does not occur only as a consequence of gene amplification. Rather, augmented levels of cyclin D1 often result from its defective regulation at the post-translational level (113).	cancers (114) and some premalignant lesions (82). CCD1 positive expression has also been reported in high-grade OSCC which appears to have a predictive value for the prognosis of the patients with lower TNM stage oral carcinoma (115). Moreover, in OSCC, increased production has been related to more aggressive tumor behavior and worse prognosis than tumors that do not overexpress CCD1 (116).
Signal Transducer and Activator of Transcription (STAT) Proteins	Members of the STAT family are latent cytoplasmic transcription factors activated by extracellular signaling proteins, such as growth factors, cytokines, hormones, and peptides. Activated STAT proteins deliver the signals by translocating into the nucleus and regulating transcription of target genes involved in normal cell functions, such as growth, apoptosis, and differentiation (117).	STAT 3 signaling is considered immunosuppressive and may protect cancer cells from recognition and lysis by cytotoxic T lymphocytes (118). Constitutive activation of STAT3 can up-regulate transcription of target genes, including cell-cycle regulators, anti-apoptotic genes, and pro-angiogenic factors, resulting in uncontrolled cellular proliferation, anti-apoptotic response, and angiogenesis, all hallmarks of cancer (119).	Both tumor and normal epithelia of SCCHN patients have shown higher levels of STAT3 expression and phosphorylated forms compared to epithelium-derived from cancer-free individuals, implicating STAT3 activation as an early step in oral carcinogenesis (120). Furthermore, highly expressed STAT3 was found in poorly differentiated OSCC tumors and has been correlated with lymph node metastasis and poor prognosis (121).

Nuclear Factor-kappa B (NF-κB)	NF-κB is a protein complex that controls DNA transcription, cytokine production, and cell survival. It acts as a ubiquitous nuclear transcription factor involved in cellular responses to stimuli such as stress, cytokines, free radicals, heavy metals, ultraviolet irradiation, and bacterial or viral antigens. It plays a key role in regulating inflammatory and immune responses (77).	In its inactive state, NF- κB is present in the cytoplasm in a complex with an inhibitory subunit IκBα which in a response to a stimulus is degraded, resulting in NF- κB being released. This activated NF- κB then translocates to the nucleus and regulates target genes, including immunoregulatory and inflammatory genes, anti-apoptotic genes, and genes that positively regulate cell proliferation. Specifically, NF-κB administers the expression of proteins with leading roles in cell proliferation, survival, immune response, and inflammation (122).	Data analysis suggests that NF-κB signaling plays an important role in oral carcinogenesis. Up-regulated NF-κB expression has been reported in OSCC, with the level increasing gradually from normal mucosa to premalignant lesions and invasive cancer, proposing that NF-κB is activated in the early stages of oral cancer development (122,123).
Activating Protein-1 (AP-1)	The AP-1 family of transcription factors consists of multiple Jun and Fos members (77).	The AP-1 complex causes multiple growth signals to converge at the transcriptional level and is involved in the regulation of cellular proliferation, differentiation, apoptosis, oncogene-induced transformation, and cancer cell invasion (124).	Constitutive activation of AP-1 binding proteins, including the Jun family and Fra-1, can be detected in oral dysplastic and OSCC cell lines. It has been associated with malignant transformation in squamous epithelial cells (125). Moreover, activation of AP-1 by transfection of c-Jun showed to induce malignant conversion to SCC in murine models (126).

4.5 Tumour suppressor genes and abnormalities in growth-inhibitory signals

Tumor suppressor genes are normal genes that slow down cell division transducing negative growth-regulatory signals, repair DNA mistakes, or lead to programmed cell death (apoptosis). Unlike oncogenes, they cause cancer when inactivated by several mechanisms, including point mutations and/or deletion, in both alleles of the gene. Although the majority of the tumor suppressor gene mutations are acquired, inherited abnormalities have been found in some family cancer syndromes, including OSCC (Table 2). Classical examples are the Rb-1 gene associated with the development of retinoblastoma and the p53 gene, connected with a wide range of neoplasms, including breast cancer and leukemia. Others are correlated with Wilms' tumor gene (WT1), neurofibromatosis, and infrequent forms of colorectal cancer. Once these genes are inactivated, the cell escapes regulated cell-cycle control, predisposing it to uncontrolled growth and division and thus contributing to malignant phenotype (127). The proteins encoded by tumor suppressor genes normally act to inhibit cell proliferation and tumor development. In many tumors, these genes are absent or inactivated, thereby removing negative regulators of cell proliferation and contributing to tumor cell proliferation. Eventually, for cancer development are needed both, the acquisition of self-sufficient growth signals through oncogene activation and the loss of growth-inhibitory signals. Growth inhibitory signals are tightly regulated by interactions of the cyclin-dependent kinases (CDK), cyclin, and the product of the retinoblastoma (*Rb*) gene. Besides, the proteins encoded by the tumor suppressor genes p16, p21, and p53 also act as inhibitors of cell-cycle progression.

Table 2. Commonly deregulated tumor suppressor gene products related to the molecular pathogenesis of OSCC.

Tumor suppressor gene products	Description	Signaling	Association with OSCC
Retinoblastoma gene (RB)	The retinoblastoma tumor suppressor (RB) regulates cell cycle progression at the restriction point between the early and late G1 phase (77).	The substantial role of the Rb pathway is evidenced by the finding of inactivation of CDKN2A in HNSCC. It has been previously shown that mutation-caused activation of CDKN2A is significantly more seldom than deletion or epigenetic switch off, which account for	Mutations in the RB pathway are happening in the OSCC onset. Loss of Rb protein expression has been reported in 66% (128) and 74 % (129) of OSCC cases, also in 64% of premalignant lesions (128).

		inactivation of the gene in up to 75% of HNSCCs (105).	
TP53	*TP53* is a tumor-suppressor gene encoding a transcription factor involved in the maintenance of genomic stability, cell cycle, DNA repair, apoptosis, and senescence (130). p53 is a major cellular stress sensor for DNA damage or oncogene activation (131). It is the most commonly mutated gene, altered in about 50% of all cancers, including 25-69% of oral cancers (132).	Mutation in TP53 occurs early in carcinogenesis and is often associated with HPV infection, due to dysregulation of p53 by the HPV E6 oncoprotein (133). Overall, the data suggest that the p53 pathway is downregulated in about 80% of HNSCCs (105).	TP53 mutation and overexpression have been correlated with poor survival in OSCC (97), (134). Positive expression of p53 protein has been seen in oral dysplastic lesions (134). Amplification of another p53 family member, TP63, is observed in about 80% of HNSCC (111).
$p21^{WAF1}$	$p21^{WAF1}$ is an influential cell-cycle protein inhibitor whose expression is activated by wild-type p53 (105).	p21 interacts with cyclin/CDK, leading to the cell-cycle arrest. Thus, p21 plays a major role in mediating the growth-suppressing, as well as the apoptosis promoting, functions of p53 (135).	It was demonstrated that p21 expression is increased in premalignant and malignant oral lesions through p53-dependent and -independent pathways, suggesting that alterations in p21 may be early events in oral carcinogenesis (136). p21 gene mutations have not been described in oral cancers (105).
$p16^{INK4a}$	p16 is a protein that delays cell division by slowing the progression of the cell cycle from the G1 phase to the S phase, thereby acting as a tumor suppressor. It is encoded by the, CDKN2A gene (128).	Genetic inactivation of the p16 gene (CDKN2A) by deletion, methylation, and point mutation has been found in approximately 50% of all human cancers, including oral malignancies (137).	Altered or loss of p16 expression has been found in 83% of oral premalignant and 60% of malignant lesions (128)(140-141). Also, studies indicate a correlation between p16 and the poor prognosis of OSCC patients (140).

NOTCH	NOTCH 1, 2,3, and 4 comprise the NOTCH family of transmembrane proteins (141).	The notch pathway promotes proliferative signaling during neurogenesis and is crucial in cell-cell communication, which involves gene regulation mechanisms that control multiple cell differentiation processes during embryonic and adult life (141).	NOTCH 1-3 was found to be present in 17% of HPV positive and 26% of HPV-negative HNSCC. Loss of NOTCH signaling was found to promote tumourigenesis in HNSCC the mechanism of which remains unclear (142). Nevertheless, genomic analysis data support a tumor-suppressive role of NOTCH1 in epithelial SCC tumorigenesis (143).

4.6 Invasion and metastases

Local invasion and a tendency for dissemination to cervical lymph nodes are characteristic of OSCC. In the cancerous state cells gain a migratory ability, invade surrounding tissues, and metastasize is an intricate process. In it, cancer cells undergo a morphological change and change from a polarized, epithelial phenotype to a fibroblast-like mesenchymal phenotype, a process known as Epithelial-Mesenchymal Transition (EMT) (144). As a result migration, invasion, and metastatic progression are favored. A key target in EMT are the cadherins. A superfamily of calcium-dependent transmembrane proteins that mediate cell-cell adhesions between normal mucosal cells maintains epithelial integrity. A distinctive characteristic of EMT is the cadherin switch, describing the process of decreased E-cadherin and increased N-cadherin expression. This switch is probably contributing to increased motility but not for morphological aberrations occurring in EMT (145). However, studies regarding the cadherin switch in oral cancer are controversial. Ukpo et al. concluded that neither E- nor N-cadherin expression may serve as a predictor for nodal or distant metastasis in oropharyngeal SCC (146). Lim et al. also reported no correlation of E-cadherin with cervical lymph node metastasis in stages I and II in tongue SCC (147). In contrast, other studies have shown that diminished or complete loss of E-cadherin expression is associated with oral cancer progression, clinicopathological parameters and low p53 levels (148), lymph node metastasis, and poor prognosis (149). The role of E-

cadherins in malignant transformation of oral leukoplakia has also been investigated (150). Given the characteristic role of cadherins in establishing and maintaining epithelial stability, their use as prognostic and predictive markers is being investigated. Integrins are cation-dependent transmembrane glycoproteins acting mediating cell-cell and cell-matrix interactions. They are involved in the maintenance of tissue integrity and regulation of cell proliferation, growth, differentiation, and migration (151). In constant, loss or reduced expression of integrins β1 and α6 β4 has been observed in OSCC, correlated to loss of basal membrane proteins and vaster in poorly differentiated lesions (152). Besides, integrin α6 β4 expression has been associated with early recurrence and metastasis based on an immunohistochemical evaluation. Up-regulated integrin αvβ6 has been found in oral leukoplakia (153), and later correlated with malignant progression (154). It has been reported that human SCC cell line growth strictly requires αvβ6 protein and its up-regulation contributes to cell migration suggesting its possible critical role in regulating oral cancer growth and invasion (155). Collective data suggest that integrins may play varied and complex roles in the progression of OSCC tumors and that these roles may depend on their subunit composition. Finally, the process of oral cancer invasion and dissemination requires proteolysis of the basement membrane by enzymes, including the matrix metalloproteinases (MMPs). MMPs are a 24-member family of zinc metalloenzymes involved in extracellular matrix remodeling. They are subdivided into different groups due to structural features and substrate affinity and include collagenases, gelatinases, stromelysins, membrane-type MMPs, and new MMPs. In a normal physiological state, MMPs migrate through tissue barriers to take part in developmental and healing processes while cancer cells use similar mechanisms to degrade the extracellular matrix favouring migration and metastasis. Elevated MMP production has been detected in OSCC and many other malignancies and correlated with aggressive tumor behavior and poor patient prognosis. Gelatinases (MMP-2 and 9), stromelysins (MMP-3, -10,-11), collagenases (MMP-1 and -13), and membrane-type MMPs (MT1-MMPs) are taught to be implicated in the invasion of oral cancer (156). Basement membrane disruption by MMP-2 and MMP-9 is believed to be the initiation of neoplasm invasion. Further, studies have disclosed that metastatic oral tumors have higher MMP-2 expression than non-metastatic ones, suggesting the predictive capacity of these metalloproteinases in tumor spread (157,158). Stromelysins are likely to maintained OSCC progression as MMP-3 was reported

significantly pronounced in the invasive front while overexpression of MMP-10 and -11 has been linked to local invasiveness (156). Collectively results point that MMPs appear to be robust and reliable indicators with an important role in the progression of OSCC.

4.7 *The role of inflammation*

Inflammation is a beneficial response activated to repair tissue injury and eliminate pathogens to recover the homeostatic state. However, if dysregulated can become prolonged (chronic) involving a progressive change of the cell types present at the affected site. The immune response comprises a series of events triggered after recognition of pathogens or tissue damage, involving cells and soluble mediators, such as cytokines of the innate and adaptive immune system. Cytokines are key players in modulation in both innate and adaptive immune responses. They present a vast group of small molecular weight proteins (5-20 kD) acting on the cells that secrete them (autocrine signaling), on nearby cells (paracrine signaling), or in some instances on distant cells (endocrine signaling). Being produced by diverse cells, including macrophages, B- and T lymphocytes, fibroblasts, mast, endothelial, and various stromal cells, they are usually secreted transiently and locally in response to a variety of stimuli (159). Cytokines comprise interleukins (ILs), chemokines, interferons (IFNs), and tumor necrosis factors (TNFs).

- The human genome encodes more than 50 interleukins and related proteins, the majority of which are secreted by helper CD4 T lymphocytes, as well as by monocytes, macrophages, and endothelial cells. Interleukins promote the development and differentiation of T and B lymphocytes, and hematopoietic cells. Common representatives are IL1-17, IL-21, IL-22, IL-24, IL-25, IL-26, IL-27, IL-35, etc. Lymphocyte-produced interleukins that mediate immune responses were classified as lymphokines.
- Chemokines are signaling proteins with a major function to induce chemotaxis to guide cell migration (chemotactic cytokines). Chemokines have conserved cysteine residues allowing them to be grouped as C-C chemokines (like monocyte chemoattractant protein (MCP-1), C-X-C chemokines (like growth-related oncogene (GRO/KC), C chemokines (lymphotactin), and CXXXC chemokines (fractalkine). Functionally are divided into homeostatic chemokines, produced constitutively and responsible for basal leukocyte migration (such as CCL14, CCL19, CCL20, CCL21, CCL25, CCL27, CXCL12,

and CXCL13) and inflammatory ones, formed under pathological conditions and actively take part in the inflammatory response (like CXCL-8, CCL2, CCL3, CCL4, CCL5, CCL11, and CXCL10) (160).

- The interferons are released in response to the presence of several viruses. Interferons are named after their ability to "interfere" with viral replication by helping the cells to eradicate pathogens. IFNs also activate immune cells, such as natural killer cells and macrophages; they increase host defense by up-regulating antigen presentation via increasing the expression of major histocompatibility complex (MHC) antigens.
- TNF superfamily unites 19 members of type II transmembrane proteins. They are expressed predominantly by immune cells and mediate cell functions including immune response and inflammation, proliferation, differentiation, apoptosis, and embryogenesis. Widely known representatives are tumor necrosis factors alpha (TNF-α), Lymphotoxin-alpha (LT-α or TNF-β), etc.

Inflammation is characterized by the interplay between pro-and anti-inflammatory cytokines. Pro-inflammatory cytokines are produced predominantly by activated macrophages and are involved in the up-regulation of inflammatory reactions. Such agents are IL-1α, IL-1β, IL-6, IL-8, and TNF-α and. The anti-inflammatory cytokines are immunoregulatory molecules that control the pro-inflammatory cytokine response. Representatives include interleukin 1 receptor antagonist (IL-1RA), IL-2 IL-4, IL-10, IL-11, IL-13, interferon-gamma (IFN-*γ*), and transforming growth factor-beta (TGF-β) (159). The oral cavity is one of the most ecologically elaborate microenvironments in the human body wherein interactions between the host and microbes determine health and disease.

Common diseases affecting oral tissues such as gingivitis, periodontitis, and OLP have confirmed the pivotal role of inflammation in the pathogenic process (161). A series of investigations have revealed a close relationship between chronic inflammation and oncogenesis, suggesting immune activation and chronic inflammation as their possible causes. The inflammatory response shares various molecular targets and signaling pathways with the process of malignant transformation, such as apoptosis, increased proliferation rate, and angiogenesis. Unresolved inflammation increases the risk for cancer by supplying bioactive molecules like cytokines, growth factors, cell survival signals to avoid apoptosis, proangiogenic factors, and extracellular matrix-modifying enzymes such as metalloproteinases. Inflammatory cells are essential components of the tumor

microenvironment and can promote cancer cell proliferation and survival, as well as invasion and metastasis (162). Depending on the tumor microenvironment, cytokines can modulate an antitumoural response, but during chronic inflammation, they can also induce cell transformation and malignancy, depending on the balance of pro-and anti-inflammatory factors, their relative concentrations, cytokine receptor expression content, and the activation state of surrounding cells (163). The role of inflammation in carcinogenesis was firstly described by Rudolf Virchow, as early as 1863. Since then, multiple studies have provided abundant evidence supporting that chronic inflammation is a pathological response that can be detrimental to the host affecting cell homeostasis and metabolic processes, inducing even genomic changes, and which can promote carcinogenesis (164). The arise of at least 20% of all cancers has been associated with infection and chronic inflammation and even those cancers that do not develop as a consequence of chronic inflammation, exhibit extensive inflammatory infiltrates with abundant cytokine expression in the tumor microenvironment. The cytokine and chemokine expression profile of the tumor microenvironment may be more relevant than its specific immune cell content. Some of these proteins were found to serve as growth and survival factors that act on premalignant cells, stimulate angiogenesis, tumour progression and metastasis, and also sustain tumour-promoting inflammation, regardless of their source (165) (Fig.10).

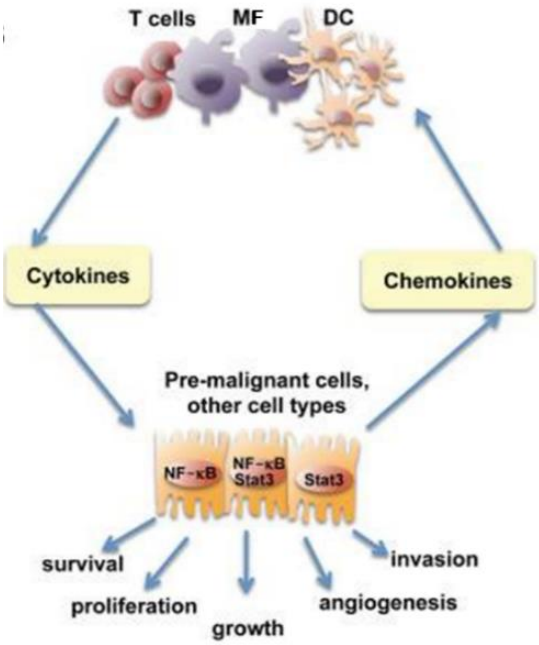


Figure 10. A major tumor-promoting mechanism is the production of cytokines by tumor-infiltrating immune cells such as T-lymphocytes (T-cells), Macrophages (MF), and Dendritic cells (DC) which leads to the activation of transcription factors such as NF-κB or STAT3 acting on premalignant cells to stimulate survival, proliferation, growth, angiogenesis, and invasion and metastasis. Source: Grivennikov et al., 2010 (165).

Through activation of downstream effectors, such as NF-κB, AP-1, and STAT transcription factors, as well as caspases, cytokines regulate the immune and inflammatory milieu to either favor anti-tumor immunity (IL-12, TRAIL, IFNγ) or increase tumor progression (IL-6, IL-17, IL-23) and also have direct effects on cancer cell growth and survival (TRAIL, FasL, TNF-α, EGFR ligands, TGF-β, IL-6) (166). Aberrant activation of NF-κB and/or STAT3 is found in over 50% of all cancers (167). The clinical significance of cytokines has grown since powerful evidence has revealed the critical involvement of NF-κB mechanism of action in carcinogenesis, apoptosis protection, and chemoresistance in neoplasia including breast, ovarian, gastric, pancreatic, and HNC (168,169). NF-κB is a key player in the inflammatory response and its signaling pathway is largely based on its role in the expression of pro-inflammatory agents including cytokines such as IL-1, IL-6, TNF-α, chemokines, and adhesion molecules (168). Their role in the pathogenesis of oral cancer and associated lesions has been long-investigated. NF-κB-dependent overproduction of certain cytokines is observed in patients with OPMD compared to individuals with pre-malignant or without oral lesions (170). Various case-control studies revealed considerably altered TNF-α, IL-1β, IL-6, and IL-8 levels detected in serum, saliva, and tissue specimens of patients with oral homogeneous and non-homogeneous leukoplakia, as well as in oral lichen planus (171–173). In vitro and in vivo studies have also demonstrated elevated expression of the same repertoire of cytokines detected in OSCC cell lines, tissues, serum, and saliva samples, demonstrating the involvement of the NF-κB pathway in the pathogenesis of oral cancer (174). Other inflammatory factors including IL-1α, IL-2, IL-10, IFN-*γ* have also shown differential expression in saliva and serum OSCC samples (169,170). The local and systemic nature of these responses suggests that altered pro-inflammatory cytokine responsiveness is tightly associated with pre-cancer and cancer and could contribute to the biological mechanism of oral carcinogenesis. Furthermore, various cytokines have been proposed as potential markers for malignant conversion risk assessment of OPMDs, early oral tumorigenesis, progression, treatment outcomes and targets for therapy (170,173,175,176).

5. *Biomarkers*

A common objective of clinicians and researchers is to noninvasively and accurately assess and monitor the physiological status of healthy and diseased people. Early disease detection can significantly influence patients' comfort, prognosis, therapeutic options, survival rates, and recurrence. Diagnosis and surveillance often require painful, invasive procedures such as biopsies and repeated blood draw. To overcome these measures, the development of a non-invasive or minimally invasive methodology would facilitate population screening for risk of disease development, initial diagnosis, and staging, as well as to monitor its progression and predict possible outcomes. The implication of biomarkers as quantifiable characteristics in large trials of major pathologies such as cancer and heart diseases is known since the 1980s (177). To date, the Food and Drug Administration (FDA) continues to promote their use in basic and clinical research, as well as to identify new potential targets as surrogates in future trials. According to the National Institute of Health (NIH), a biomarker is an objectively measured and evaluated indicator of normal biologic and pathogenic processes or pharmacologic responses to therapeutic treatment that exists in a variety of forms including antibodies, microbes, DNA, RNA, lipids, metabolites, and proteins. Alterations in their concentration, structure, function, or action can be related to the initiation, progression, or even regression of a particular disorder (178). A cancer biomarker refers to a substance or process that is indicative for the presence of malignant formation in the body. A valid biomarker for malignant tumors requires to be specific for the tumor type, detectable in a high level in the patient and undetectable or present in low level in not affected people and easily quantifiable in a clinical sample. In terms of clinical utility, regarding their function, cancer biomarkers can be classified as diagnostic, prognostic, predictive, and therapeutic. A diagnostic biomarker should be directly correlated with the presence of the disease and the most specific and sensitive at its early stage. A predictive biomarker is useful to indicate the treatment response and thus defining subpopulations of patients that are likely going to benefit from a specific therapy. A clinical or biological parameter that gives information about the possible course of the disease and the patient survival is defined as prognostic biomarkers. A therapeutic one is generally a substance that could be used as a target for therapy (179). The clinical implementation of any single or a panel of biomarkers for health risk assessment or

prognosis requires an extensively controlled multiphase process preceded by discovery and validation phases. Potential candidates are subjected to comprehensive evaluation, including preclinical and academic verification, before FDA assessment and approval. Currently, there are only 24 biomarkers approved by the American FDA applied in clinical practice and associated with some cancer types. They comprise glycosylated proteins and genetic hotspots (DNA fragments likely to mutate). Some of them are human epidermal growth factor receptor 2 (HER2/neu) and Cancer antigen 15-3 and 27-29 (CA15-3 or MUC1) (CA 27-29) for breast cancer, prostate-specific antigen (PSA or Pro2PSA) for prostate cancer, cancer antigen 125 (CA125 or Muc16) and human epididymis protein 4 (HE4) for ovarian cancer, α-Fetoprotein (AFP) for hepatic cancer, Thyroglobulin (Tg) for thyroid cancer, etc. The restricted number of approved biomarkers is ascribed to the highly heterogenic nature of the cancer cells, even of the same origin (180,181).

5.1 *Biomarker discovery in OSCC*

Despite improvements in therapeutic strategies, OSCC survival rates have not ameliorated over the past years mostly because of late diagnosis and frequent recurrences. The challenge of this decade, therefore, is to reduce both mortality and morbidity of this disease through the identification of potential biomarkers for early detection, monitoring of disease progression, and recognition of therapeutic targets. A special emphasis is given to the early diagnosis, being the ultimate goal of oral cancer prevention and considered to have a positive impact on the successful treatment, prognosis, and quality of life of these patients. Moreover, the early indication of oral mucosal lesions that can precede cancer such as oral leukoplakia/ erythroplakia is critical for prognosis and survival rates. Existing studies suggest that it is unlikely a single biomarker to provide the specificity and sensitivity necessary for the identification of a certain tumor type. In this regard, the focus of the latest research is the combination of a number of individual analytes in a broad panel thought to have more reasonable diagnostic accuracy. However, to assemble and validate a panel of biomarkers is a huge challenge requiring integration and complex analysis from multiple experimental sources using interdisciplinary tools, the so-called ‘’-omics’’ technologies (182). They refer to an advanced and powerful group of high-throughput research tools, such as genomics, transcriptomics, proteomics, metabolomics, glycomics, lipidomics (Fig.11). These tools are based on the comprehensive analysis of genetic

information, including information from DNA, RNA, proteins, and metabolites from tissue samples, cell lines, and body fluids (183,184). Emerging OSCC data hold great promise to overcome the goal for the discovery of highly specific diagnostic, predictive and prognostic molecular biomarkers.

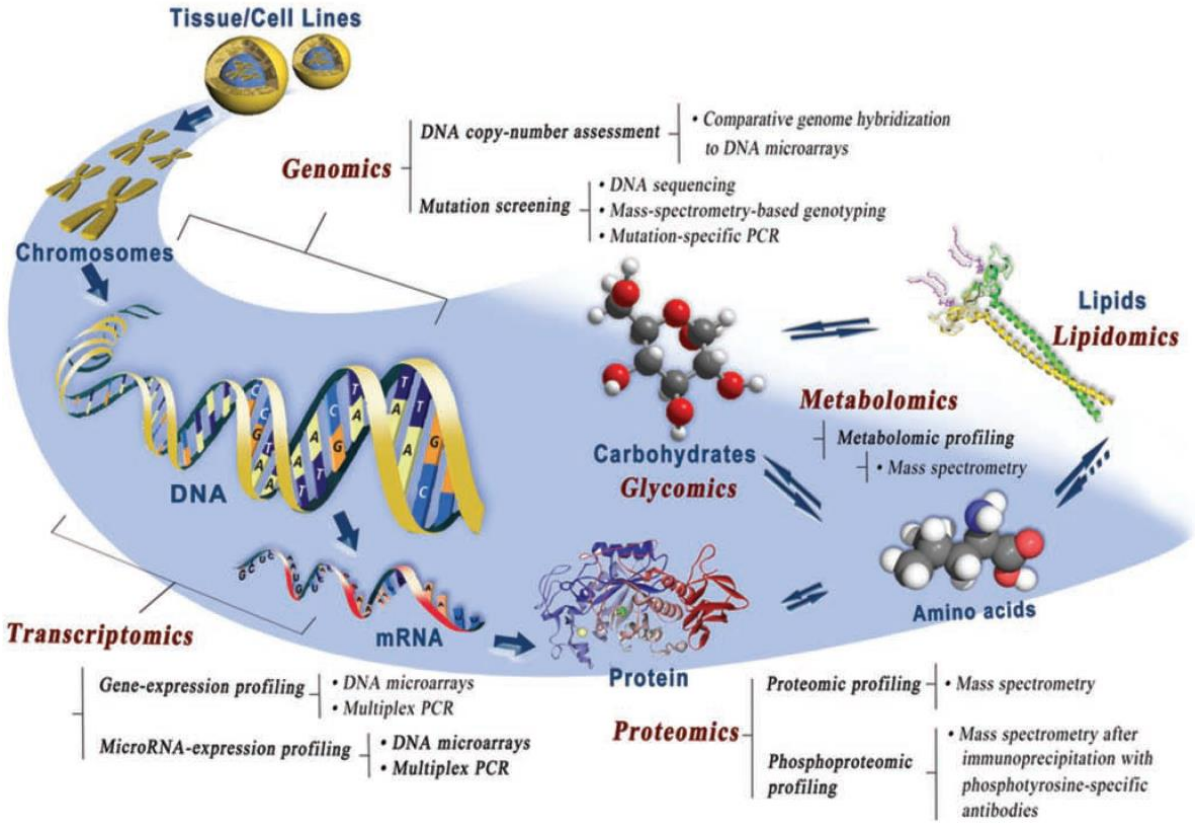


Figure 11. A scheme of ‘’-omics’’ technologies, their corresponding targets for analysis, and assessment methods. DNA (genomics) is first transcribed to mRNA (transcriptomics) and translated into protein (proteomics) which can catalyze reactions that act on and give rise to metabolites (metabolomics), glycoproteins and carbohydrates (glycomics), and various lipids (lipidomics) Sources Wu R.Q. et al. (184).

To generate an –omic expression profile, specimen collection is required. Tumour tissue and body fluid such as saliva or blood (serum and plasma) can potentially carry whole cells also protein, DNA, and RNA species allowing detection of cancer-related cellular changes. Formalin-fixed paraffin-embedded tissue (FFPET) samples, fresh frozen tissues as well as in vitro models of tumor-derived cell lines have been utilized in OSCC studies. Physiological variations among human tissues and sample unavailability can present important limitations in the comparative analysis of normal and cancerous mucosa. Compared to the challenging diagnostic utility of tissue biopsy, body fluids have gained much more attention for biomarker identification. Body fluids-derived biomarkers showed notable advantages

over the conventional histopathological method, offering non-invasive testing with reduced risk for the patients, cost, and diagnosis time (185). Blood serum and plasma are applied in almost all types of cancer research due to the content of multiple measurable molecular elements in the form of circulating cells, proteins, peptides, metabolites, and cell-free DNA and RNA. Besides, biomarkers in saliva and/or plasma can reveal carcinogenic processes and can be used to monitor its progression/remission (186). In the past decades, the diagnostic properties of saliva have been repeatedly proven to contribute to the evolvement of the complex science of salivomics. As a subset of the larger field of molecular diagnostics, yet has been recognized as a central player in a variety of biomedical basic and clinical areas (187). Besides, it is pre-programmed to have a certain composition in response to events in an oral cavity, therefore the first biomarker for breast cancer is HER2/neu, found in saliva (188). Saliva contains water (95%), proteins, minerals, nucleic acids, and electrolytes at detectable and quantifiable levels that can help to discriminate health from disease (186). Its utilization for disease screening and diagnosis is advantageous over other body fluids because of the easy accessibility, inexpensiveness, not invasive collecting procedure, and multiple sampling for monitoring the disease development. The direct contact of saliva with oral lesions has turned into a highly desirable choice for studying oral disorders, and particularly OSCC (189–191). A novel focus of research is the implication of salivary diagnostics for the early detection of oral cancers.

5.2 *Salivary biomarkers for Oral Cancer detection*

Proteins are key compounds in biosynthesis, cell, tissue, and organ signaling and provide cell and tissue structural stability in living organisms. They appeared to be attractive targets for biomarker discovery particularly those which are regulatory molecules in relevant cell signaling pathways. Protein molecular biomarkers are particularly popular due to the availability of a large range of analytical instrumentation, which can identify and quantify proteins in complex biological samples (192). Head and neck cancer salivary proteome is differs quantitatively from the one of normal, healthy saliva (193). Hence, the identification of proteins is either fractionated by gel electrophoresis or digestion by enzymatic procedures to produce peptide mixtures involving high-throughput technologies enabling the assessment of cellular whole protein complements in tissues or secreted proteins in biological fluids (194). So far, about 3000 proteins have been identified in saliva by using

various analytical platforms, advances in mass spectrometry, and a combination of data from multiple groups (195). OSCC associated studies have demonstrated aberrant expression of proteins related to cell metabolism and structure, adhesion and motility, signal transduction, inflammation, including oncoproteins (196). To meet the clinical requirement for high accuracy in the detection of OSCC, a strong board of specific and sensitive candidates is required. Some potential biomarkers identified from the saliva of OSCC patients are listed in Figure 12. Katakura *et al.* (197), in 2007, examined 20 healthy patients and 19 patients with oral cancer and estimated salivary levels of cytokines (IL-6, IL-8, IL-β1), which showed increased concentration in cancer as compared to the control cases, suggested saliva as an important screening tool. CD44 is elevated in the majority of head and neck squamous cell carcinoma (HNSCC) and distinguishes cancer from benign diseases with high specificity (198). In the saliva of OSCC patients, three known markers were found to be four-fold increased, such as cytokeratin 19 fragment (Cyfra21–1), cancer antigen 125 (CA-125), and tissue polypeptide antigen (TPS) (199). Fibrin, transferrin, Ig-heavy chain constant region gamma, cofilin-1 (200), salivary endothelial levels (201), pro-inflammatory cytokine IL-6, TNF-α, and antibodies responsive to gene aberrations, such as anti-p53 antibodies, were also identified (202). The latest techniques have been followed by Gallo *et al.* (203) for the identification of OSCC proteomic signatures. The group suggested a predictive model and analysed it through mass spectrometry of saliva from 45 OSCC patients and 30 control subjects to investigate the diagnostic and prognostic potential of the salivary proteome. Different neural networks for prognostic and diagnostic accuracy were used which indicated that selecting a particular predictive model remains under investigation. The systemic analysis of salivary proteomic biomarkers is rapidly advancing, offering an attractive screening tool to turn salivary diagnostics into clinical and commercial reality to combat oral cancer.

Molecule	Saliva
	Proteomics*
Albumin	●[a]
Cancer antigen 125 (CA125)	●[b]
Catalase	●[d]
CD44	●[d]
CD59	●[d]
Cofilin-1	●[e]
Cyfra 21-1	●[b]
Endothelin-1	●[g]
Glutathione	●[h]
Interleukin 1α (IL-1α)	●[c]
Interleukin 1β (IL-1β)	●[c]
Interleukin 6 (IL-6)	●[c]
Interleukin 8 (IL-8)	●[c]
Immunoglobulin heavy chain constant region gamma	●[e]
Mac-2 binding protein (M2BP)	●[c]
MRP14	●[d]
p53 antibodies	●[c]
Profilin	●[d]
S100 calcium binding protein	●[e]
Telomerase	●[i]
TNF-α	●[c]
Tissue polypeptide antigen (TPA)	●[c,f]
Transferrin	●[e]
Transthyretin	●[e]
α-Amylase	●[a]
β Fibrin	●[e]

[a] Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS).
[b] Immunoradiometric assay.
[c] Enzyme-linked immunosorbent assay (ELISA).
[d] Immunoblot.
[e] Liquid chromatograph/mass spectrometer (LC/MS).
[f] Time-resolved fluoroimmunoassay (TRFIA).
[g] Quantitative real time RT-PCR.
[h] High performance liquid chromatography (HPLC).
[i] PCR and ELISA.
[j] DNA microarray.
[k] Methylation-specific PCR.
[l] RT-PCR and RNA microarray.
[m] Real-time quantitative polymerase chain reaction and RNA microarray.

Figure 12. Potential salivary biomarkers for detection of OSCC. Adapted from Jia-Yo Wu et al. (191)

Changes in protein posttranslational modifications (PTMs) have an important role in studying disease etiology and progression. It has been shown that defective glycosylation accompanies many chronic and infectious conditions and is a common feature of tumor cells that may affect N- and O-glycans on glycoproteins, glycolipids, or glycosaminoglycans (204). Among the most common N-linked glycans expressed in mammalian glycoproteins are oligomannose (high-mannose), hybrid and complex bi-, tri-, and tetrantennary types (Fig.13). Glycan biosynthesis is a complex and tightly regulated process involving a network of hundreds of genes. Unlike DNA and proteins, the biosynthetic pathway is not template-driven, resulting in complicated and variable glycosylation profiles even for single glycoproteins. Their heterogeneity, in terms of stereochemistry, makes structural analysis a difficult task, with relatively recent technologies enabling the expansion of glycan investigation (205).

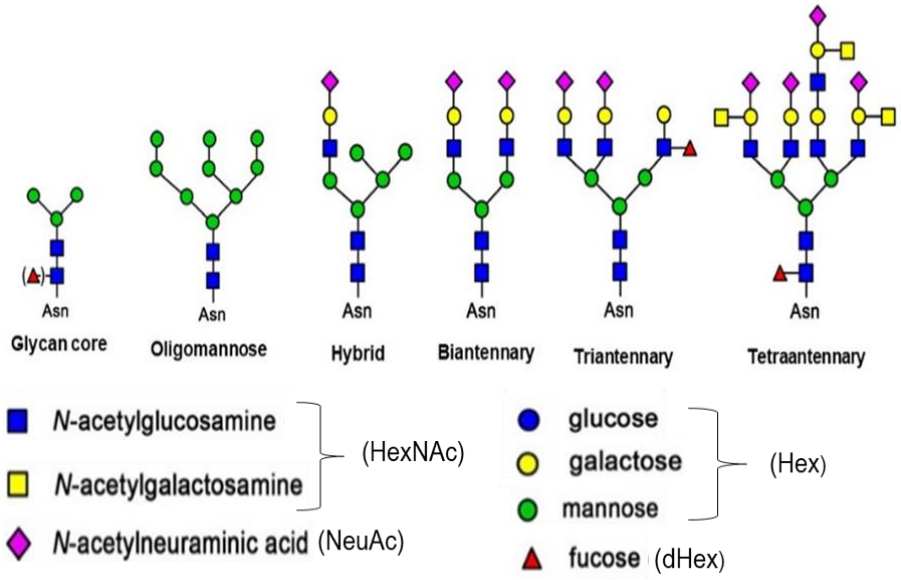


Figure 13. Commonly expressed *N*-linked glycans on human glycoproteins. N-acetylglucosamine and N-acetylgalactosamine (HexNac), glucose, galactose and mannose (Hex), Fucose (dHex), and Sialic Acid (NeuAc). *N*-glycans can form high-mannose, hybrid or complex types of structures. Complex types may present bi, tri-, and tetra- antennary configurations. Source: Essentials in Glycobiology, 3rd edition, 2017 (206).

In the past decade research in the glycomics field gave insight into the biological importance of the serum *N*-glycome in human health and disease. Special emphasis was placed on exploring the connection between altered *N*-glycosylation of glycoproteins and different diseases, particularly in the study of cancer (207). The main methods of glycosylation analysis involve the separation of released glycans by HILIC (hydrophilic-interaction chromatography), UHPLC, CE (capillary electrophoresis), lectin affinity, and MS. Typically HPLC based glycan analysis involves the removal of glycans from glycoproteins by enzymatic digestion. Glycans coupled to a protein through the nitrogen atom (*N*-linked) of asparagine (Asn) side chains can be released by amidase action by PNGaseF (peptide *N*-glycosidase F). *O*-glycans are conjugated to serine (Ser) or threonine (Thr) without a clear consensus motif and often are cleaved by a chemical reaction due to the absence of a universal enzyme. The utilization of the UHPLC coupled MS platform for analysis is robust, reproducible, has a high dynamic range, and is quantitative, which are fundamental requirements in the discovery and validation of clinical markers. Potential biomarkers derived from serum glycoproteins for liver, pancreatic, prostate, ovarian, breast, lung, and stomach cancers have been proposed (207). Furthermore, a recent study has investigated

serum *N*-glycomes and anti-carbohydrate antibodies from normal populations and OSCC patients and suggests aberrant glycan structures and anti-carbohydrate antibody profiling in OSCC patients as diagnostic biomarkers (208). In addition, it was stated that acetylgalactosaminyltransferase 2 (GALNT2), found overexpressed in 73% of the examined OSCC tissue specimens is enhancing cancer migration and invasion by regulating EGFR glycosylation and activity (209). It is known that saliva comprises high levels of glycoproteins including mucins, salivary agglutinin, secretory immunoglobulins (IgA), lactoferrin, amylase, and proline-rich glycoproteins. A few investigations of the intricate and highly abundant glycan population in human saliva has been described reporting different number of identified sugars. Depending on the sample size and saliva type, published studies disclose from 60 *N*-glycans in parotid saliva to 78 and 265 in whole non-stimulated saliva. Studies of oral cancer associated with aberrant protein *N*-glycosylation are deficient. However, a preliminary study showed that individuals with OC display a differential *N*-glycosylation pattern compared to controls, based on salivary *N*-glycome analysis (210). Glycomics methods are less developed than proteomics partly because of the inherent challenges associated with their analysis. Despite the advances in glycan analysis, glycan salivary testing is still at the initial stage. Identifying which N-glycans contribute most to the carcinogenic process, may lead to their use in the diagnosis, prognosis, and even treatment strategies of malignant diseases, including OSCC.

II. *OBJECTIVES*

Early diagnosis has become an ultimate goal in the management of oral cancer after prevention. The discovery of non-invasive biomarkers, such as those from saliva could facilitate early disease detection or provide important information for risk stratification, prediction, and prognosis. To deliver new diagnostic tools, an investigation has been undertaken to determine novel biomarkers at a molecular level using comparative protein and glycan expression levels between cancer, pre-cancer, and normal tissue conditions. Applying salivary testing, the current basic research aims to identify biomolecule/s that can serve as a specific and reliable screening tool for the diagnosis of patients with OSCC and associated lesions.

The general objective is to analyse salivary inflammatory cytokines, as well as salivary proteome and glycan profiles to identify potential biomarkers for early OSCC diagnosis.

Specific objectives:

1. To evaluate salivary levels of a panel of inflammatory cytokines, including IL-1α, IL-6, IL-8, TNF-α, IP-10, MCP-1, PF-4, and HCC-1 in the following groups of subjects:

- patients with homogeneous leukoplakia (HL)
- patients with proliferative verrucous leukoplakia (PVL)
- patients diagnosed at early (I&II) and advanced (III &IV) OSCC stages
- individuals without oral lesions (controls)

1a. To estimate whether there are quantitative differences in the levels of the target cytokines between the groups. Especially, if there are notable differences between the OPMD and OSCC groups.

1b. To determine whether altered cytokine expression in any of the four pathology groups (HL, PVL, early, and advanced OSCC) shows a greater association with some of the groups.

1c. To assess, whether significantly aberrant cytokine expressions correlate with any of the patients' clinical parameters (age, sex, smoking habits, lesion location, clinical type, histologic features).

1d. To evaluate if any of the target inflammatory cytokines can be considered as a potential diagnostic biomarker for the studied pathologies.

2. To characterise salivary *N*-glycome profiles of:

- controls
- patients with PVL
- patients at early stages (I&II) of OSCC
- patients at advanced stages (II&III) of OSCC

2a. To compare whether there are quantitative differences between the *N*-glycan profiles of the selected groups.

3. To construct a whole saliva protein library in order to determine proteome profiles of

- controls
- patients with PVL
- patients at early stages (I&II) of OSCC
- patients at advanced stages (II&III) of OSCC

3a. To assess whether there are quantitative differences between the proteome profiles of the studied groups and to outline potential biomarkers for early diagnosis of OSCC and associated lesions.

III. *MATERIALS AND METHODS*

This investigation is part of a larger study funded under the TRACT project by Marie Skłodowska-Curie Actions (MSCA) - H2020 Framework Programme to identify salivary biomarkers of inflammation, whole saliva proteome, and *N*-glycome profiles in an established and well-characterized cohort of volunteers. The study has been completed within three phases. The initial phase included participants' sample collection during clinical sessions and adaptation of laboratory protocols for saliva processing and further molecular analysis. Experimental work and acquired data analysis were carried out during the second stage. Study outcomes and conclusions were drafted as anticipated publications in the final, third phase.

1. *Participants*

The study was authorized by the medical ethics Review Board of the General University Hospital of Valencia (HGUV) with approval № H1480794580696 in compliance with the Helsinki Declaration. Patients and control individuals were selected at the Service of Stomatology and Maxillofacial Surgery (HGUV) and the Dental clinic of the University of Valencia, in the period between 2017 and 2019. Following the established case-control inclusion criteria, a total of 157 subjects were enrolled in the study, providing a signed informed consent and voluntarily donating biological samples for scientific purposes. Saliva specimens were distributed into three groups:

- **Group 1** consisted of 66 samples of patients with oral lesions referred to as OPMDs among which 33 were diagnosed with homogeneous leukoplakia (HL) and 33 with proliferative verrucous leukoplakia (PVL).
- **Group 2** was composed of 66 samples of OSCC patients among which 33 were diagnosed at early (I and II) and 33 at advanced (III and IV) clinical stages.
- **Group 3** comprised 25 oral fluid specimens of healthy individuals without visible oral lesions, assigned as Controls.

Inclusion criteria of Group 1

- Patient with clinically diagnosed Leukoplakia (both HL and PVL) in accordance to the Van der Waal criteria (211) and based on the clinical histopathological report from lesion/s incisional biopsy.
- Patients who have not been associated with pharmacological and/or surgical treatment.
- Without any salivary gland disorders.
- Patients aged more than > 45 years.

Inclusion criteria of Group 2

- Patients with clinically diagnosed SCC in the oral cavity based on histopathological analysis from an incisional biopsy. All tumors were graded, according to the TNM (Tumour Node Metastasis) classification system (46)
- Patients who have not previously undergone surgical intervention and/or therapy treatment (radio and/or chemotherapy).
- Without salivary gland disorders.
- Patients aged more than > 45 years.

Inclusion criteria of Group 3

- Individuals, age- and sex-matched to groups 1 and 2.
- Without present oral and/or salivary gland disorders.
- Without any acute

1.1 *Samples*

All the participants were advised to refrain from eating, drinking, smoking, and oral hygiene procedures for at least 1-hour before sampling. Samples were obtained after clinically confirmed diagnosis and before any further treatment. The subjects were asked to expectorate saliva in 15 ml sterile tubes equipped with a small funnel for convenience within 5 min. On average between 1-3 ml of unstimulated whole saliva was collected from each person (212). Some patients were sampled twice to obtain a higher saliva yield. Samples with visible traces from blood were discarded from the study.

1.2 *Saliva processing and storage*

Collected saliva samples were immediately centrifuged in at 3000 rpm at 4°C for 15 min, to obtain a clear supernatant, devoid of any food particles and debris (212). Clarified saliva supernatant was collected and aliquoted into 0.5 ml tubes for immunoassay. For proteomic and glycomic studies the total content (supernatant & pellet) was saved unseparated. All the samples were placed at -20°C for 2 -4 h prior to the long-term preservation in -80 °C freezers until further processing.

1.3 *Multiplex salivary cytokine analysis*

1.3.1 *Luminex xMAP technology*

Saliva contains proteins with a very large dynamic range (213). It is therefore challenging to detect low abundance proteins such as cytokines and chemokines in the presence of high abundance proteins (e.g. amylases, proline-rich proteins, statherin, histatin, mucin, and cystatins, etc.) (195). For quantitative protein detection, conventional enzyme-linked immunosorbent assay (ELISA) has been a gold standard method since first introduced in the early 1970s. Utilizing two distinct epitope binding antibodies "sandwiched" between the antigen, ELISAs achieve protein quantitation with specificity and sensitivity. Building upon the 'sandwich' ELISA concept, multiplex antibody-based strategies have been developed as alternatives to ELISA. The most common and well-established format for such assays utilizes antibody-conjugated microspheres (beads) from Luminex Corporation. Luminex 200 ™ platform (Fig.14) for biomarker screening and protein analysis consists of magnetic bead-based immunoassays, a detection instrument, and software.

- Bead-based immunoassay

  MILLIPLEX® multiplex assays include analyte-specific capture antibodies conjugated to xMAP® beads. MILLIPLEX® assays are analytically validated for sensitivity, specificity, reproducibility, and wide dynamic range.
- Luminex® Analyzer

  Luminex 200 ™ is a flexible analyzer based on the principles of flow cytometry, which integrates xMAP detection components, such as lasers, optics, advanced fluidics, and a high-speed digital signal processor.
- Software

XPONENT® software package designed for protocol-based data acquisition with robust data regression analysis.

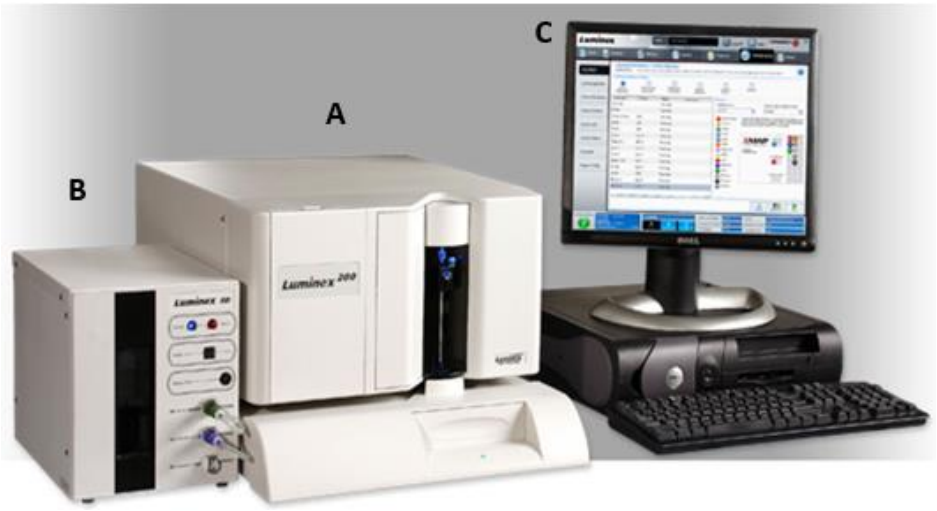


Figure 14. Luminex ® 200 System™ (Luminex Corporation, Austin, TX). The total system includes the Luminex 200 analyzer and plate handling platform (A), the Luminex SD™ sheath fluid delivery system(B), and a PC with specific software (C).

Luminex multianalyte profiling (xMAP) technology has emerged as a useful platform for simultaneous detection and quantification of multiple targets in a single, low volume sample, allowing real-time tracking of antibody-antigen interactions within an increased detection range (214). Millipore™ multiplex immunoassays utilize Luminex xMAP technology to combine the efficiency of multiplexing with the accuracy, sensitivity, reproducibility, and simplicity of ELISA (Fig.15).

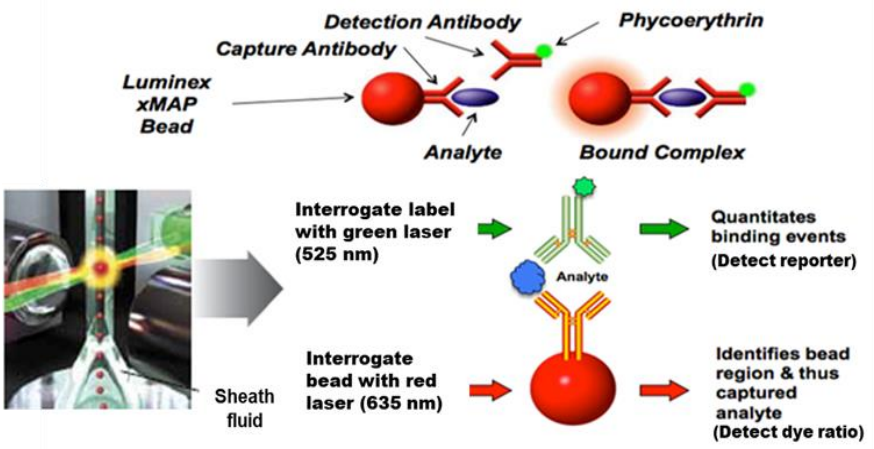


Figure 15. A general overview of the Luminex xMAP detection scheme. The assays are based upon the use of magnetic beads that have been internally dyed with red and infrared fluorophores of differing intensities. Each bead is given a unique number allowing differentiation. Individual bead sets are coated with a capture antibody qualified for one specific analyte of interest. The captured molecule from a sample is detected using an analyte-specific biotinylated antibody that binds to the appropriate epitope of the immobilized complex, plus streptavidin-conjugated R-phycoerythrin (SA-RPE). The use of different coloured beads enables simultaneous detection of multiple proteins in the same sample. Upon completion of the sandwich immunoassay, a dual detection flow cytometry allows precision fluidics (sheath fluid) to align the beads - bound complexes in a single file through a flow cell where two lasers excite the beads individually. The red laser (635 nm) excites the dyes in each bead, identifying its spectral address. The green laser (525 nm) excites the reporter molecule associated with the bead, allowing quantitation of the captured analyte. Adapted after Thermo Fisher Scientific Inc. (https://www.thermofisher.com/blog/behindthebench/luminex-bead-based-immunoassays-drive-immunoassays-towards-higher-content-biomarker-discovery).

1.3.2 *Salivary cytokine analysis*

Multiplexed immunoassays were carried out at the Molecular Oncology laboratory (University General Hospital of Valencia) utilizing the Luminex 200 ™ platform. Salivary cytokine levels were detected and quantified in 33 control, 33 homogeneous leukoplakia, 33 Proliferative verrucous leukoplakia, 33 early (I &II), and 33 advanced (III&IV) oral squamous cell carcinoma saliva samples. According to the target proteins, three Millipore's Milliplex Human Cytokine Assay kits were used including "6-plex" Panel of pro-and anti-inflammatory IL-1α, IL-6, IL-8, IP-10, MCP-1, and TNF-α (Cat.# HCYTOMAG-60K), HCC-1 (Cat.# HCTP3MAG-63K), and PF-4 (Cat.# HCVD3MAG-67K), as per manufacturer's protocols (6). Reagents provided in these kits included magnetic beads, monoclonal antibodies, standards, quality controls, assay diluents, secondary antibodies, biotin diluents, streptavidin conjugated to the fluorescent protein, R-phycoerythrin (streptavidin-RPE), washing buffer concentrates as well as the 96-well filter plates. Saliva samples were thawed directly on the day of analysis. Contents were mixed gently with a Pasteur pipette, centrifuged at 1500 rpm at 4°C for 15 min, and diluted in a 1:2 ratio to reduce viscosity. Working solutions were prepared daily. Protein standards and quality controls were prepared, within one hour prior to the beginning of the assays, by reconstituting in assay diluent and performing serial dilutions according to manufacturer specifications. Data from the reactions were acquired using Luminex 200 ™ plate reader, while a digital processor

managed data output and the XPONENT™ software returned data as Median Fluorescence Intensity (MFI) and concentration (pg/mL).

1.3.3 *Statistical analysis*

To evaluate differences in inflammatory proteins levels between control individuals, HL, PVL, early and advanced OSCC patients (objective 1a), the comparative analysis had been performed as follows:

- Determination of Level of Quantification (LOQ), relevant for concentration data.
- Multidimensional analysis to determine outliers (principal components to analyse variability between samples).
- Determination of coefficient of variation between the studied groups.
- Intra- and inter- groups Chi-square test for each analyte.
- Mann-Whitney and Kruskal-Wallis followed by Dunn's multiple comparisons non-parametric tests for each analyte.
- Differential LOQ enrichment analysis of the pathologies groups compared to the control one based on point dispersion with false-positive Benjamini-Hochberg correction (FDR).

To determine if altered cytokine expression showed a greater association with any of the pathology groups (objective 1b), the subsequent analysis were carried out:

- Comparative analysis based on the mean values of significant analytes for each group: Boxplots, Heatmaps, Venn Diagrams, Histograms.
- Vulcan plot (Plots XY) of base 2 logarithmically transformed (Log2) LOQ versus FD of each cytokine in pair-wise group comparisons.
- Regression and/or correlation analysis between groups and significant analytes.
- Cluster analysis of the cytokine average concentration in the five groups.

To determine whether there is an association between altered cytokine levels and patients clinical parameters including lesion's location, clinical type, histological features, etc. (objective 1c) the following analysis have been conducted:

- Data indexing of significantly altered cytokines and their clinical metadata in a Bayesian network probabilistic model over a set of variables (LOQ values, clinical parameters, and pathologies).
- The conditional probability of the relationship of each cytokine to belong to one of the four pathological groups, based on the LOQ value range and the probability of different clinical observations.

To achieve objective 1d:

- The Bayesian network model has been applied to predict the critical LOQ of each cytokine so that one or more analytes are determinants as a marker or risk markers related to one of the four pathology groups, given the presence or not of other specific clinical phenotypes.

1.4 *Glycomic analysis*

The main methods of glycosylation analysis involve *N*-glycan release from proteins, UHPLC (ultra-high-performance liquid chromatography) separation, and MS (mass spectrometry) detection and quantification (Fig.16) Typically UHPLC based analysis initiates with the cleavage of oligosaccharides from glycoproteins by enzymatic digestion by PNGaseF (peptide *N*-glycosidase F amidase), labeling with a fluorescent tag (2-aminobenzamine) and subsequent separation using chromatographic methods. For the determination of structural information, MS has demonstrated to be a powerful tool where neutral glycan fragmentation yields singly and doubly charged ions analyzed in a negative mode for providing a stronger signal. Acidic glycan gives ions in a higher charge state due to anionic groups.

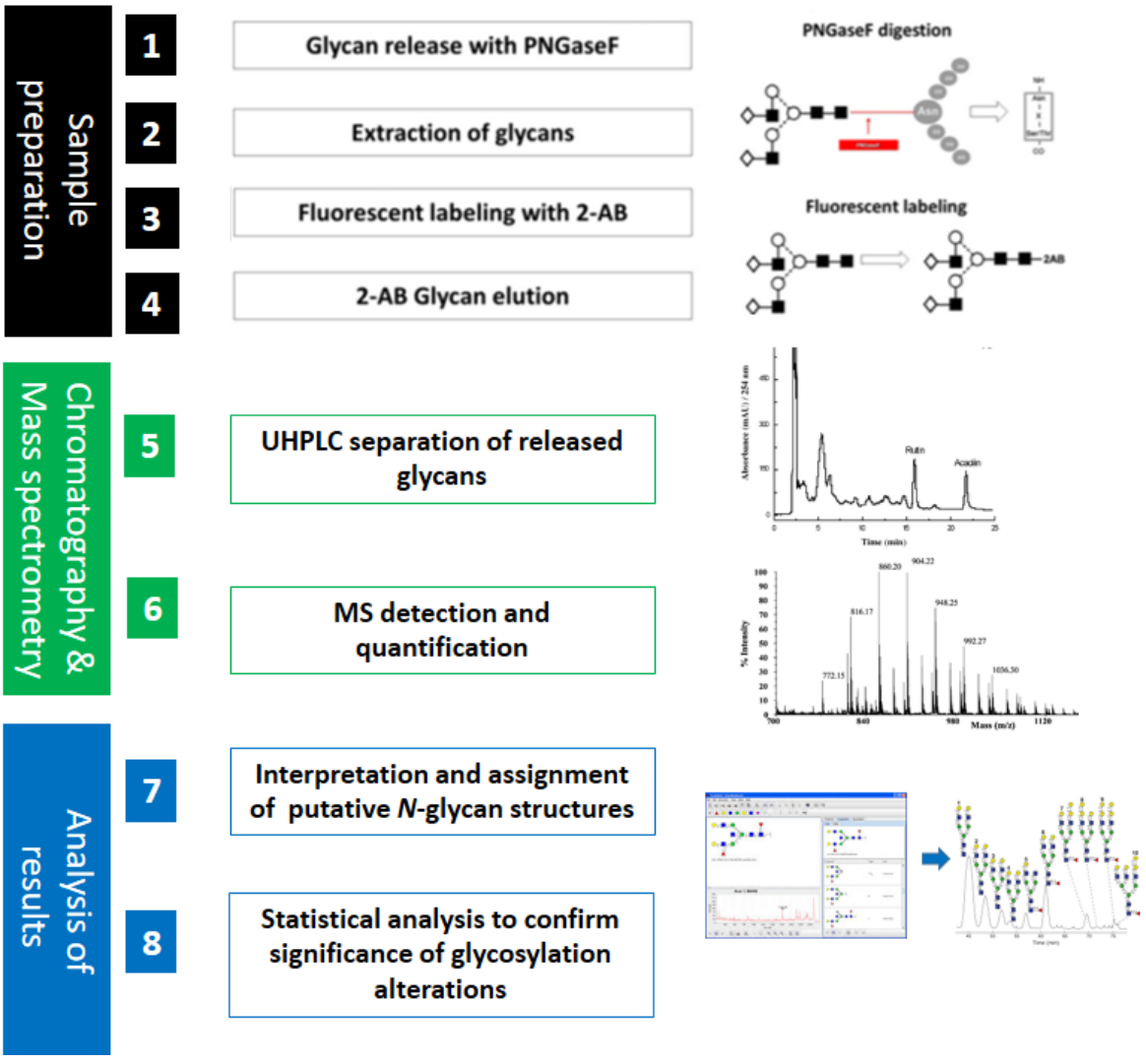


Figure 16. Technology platform for multidimensional glycan analysis and high throughput profiling. Enzymatically released salivary *N*-glycans were labeled with 2AB and profiled using HILIC separation

by UHPLC. MS analysis provides spectral peak data with theoretical glycan mass values, as well as relative quantitation of the identified fragments. Further structural annotation of glycan MS data is computer-assisted by the GlycoWorkbench software, an online tool for glycoform analysis. It performs a database search of defined theoretical masses, calculates expected glycan fragmentation and relative m/z values with the most probable identity of all compositions tested and emits a putative monosaccharide composition. Adapted after Adamczyk et al. (207)

1.4.1 *Saliva pre-analytical processing*

For *N*-glycan profiling, 20 out of the 165 saliva samples were selected and equally distributed into four homogeneous groups: 5 Controls, 5 PVL, 5 early (I and II), and 5 advanced (III and IV) OSCC stages, respectively. After one freeze-thaw cycle and centrifugation within the described above conditions, saliva was subjected to protein quantification. The total protein concentration of individual samples was determined by the Bradford assay (215) using the Bio-Rad QS Protein Assay (Bio-Rad, Hercules, CA, USA) and evaluated by Perkin Elmer spectrophotometer. 250 µg of protein from each of the five samples belonging to the same group were pooled together. Approximately 1 mL mixture of clarified saliva was precipitated with an equal amount of ice-cold 30% TCA (trichloroacetic acid) (Sigma Aldrich) for 10 min and centrifuged for 5 min at 13,200 rpm at 4°C. Supernatants were discarded and TCA residuals were removed. Pellets were re-suspended in ice-cold acetone (200 µl) and centrifuged for 1 min at 13,200 rpm at 4°C. Acetone was discarded and the pellets were air-dried.

1.4.2 *N-glycan release and fluorescent labeling*

For technical repetition, three independent aliquots corresponding to 1, 250 mg of protein pellet from each group pooled mix were homogenized in 200 mM sodium bicarbonate buffer pH 7 using Sonoplus sonicator in 10 s impulses at 20 Hz for a total of 30 s. Mixtures were transferred to 10 kDa molecular weight cut-off (MWCO) filter units for concentration and purification. After a centrifugation cycle, NaHCO3 was discarded and protein sample volume was reduced to 100 µL. To enable protein relaxation and accessibility of potential shielded glycosylation sites 8M urea in 100 mM Tris, pH 8,5 was applied. Further, protein reduction and alkylation were achieved with 10 mM DTT (dithiothreitol) and 55 mM IAA (iodoacetamide), respectively, and finally buffer-exchanged to 50mm ABC (ammonium bicarbonate). For *N*-glycan release, the resulted pellets from each aliquot were subjected to enzymatic digestion (deglycosylation) with 3U glycerol-free PNGase F (*Flavobacterium*

meningosepticum, New England Biolabs) overnight (18 h) at 37°C with agitation (216). Released glycans were eluted from a spin filter, while retained material was washed twice with water; eluates were pooled and reduced to dryness via vacuum centrifugation. Dry samples were re-suspended in 50 µL of 1% formic acid for the conversion of glycosylamines to reducing sugars. Fluorescent labeling of released glycans was carried out by reductive amination with 2-AB (aminobenzamide) in the presence of sodium cyanoborohydride in 30:70 Acetic acid/ DMSO (dimethyl sulfoxide) solution for 2h at 65°C. Removal of the excess 2-AB label was performed as reported previously. Collected fractions were reduced to dryness via vacuum centrifugation and stored at -30 °C for further analysis (216).

1.4.3 *LC-FLR-MS analysis*

Online coupled fluorescence (FLR)-Mass spectrometry detection was carried out using Vanquish™ Horizon UHPLC (Thermo Scientific, Germering, Germany) and a Q Exactive™ Plus Hybrid Orbitrap MS instrument with BioPharma Option equipped with an Ion Max source with a HESI-II probe (Thermo Scientific, Bremen, Germany). All data were acquired using Thermo Scientific™ Xcalibur™ software 4.0. UHPLC was equipped with a Waters BEH Glycan amide column (150×1.0 mm ID, 1.7 µm). 2-AB derivatized *N*-glycans were reconstituted in 80% acetonitrile and injection volume was 10 µl. Injections for each sample were performed in triplicate. Flow rate, 0.15 ml/min and column temperature were maintained at 60 °C; Eluent A is ammonium formate in water (pH 4.4), B is acetonitrile delivering the following binary gradient: 72% B was held for one minute, followed by a linear gradient to 57% B over 30 min. The column washing step was carried out at 30% B for 4 minutes and initial conditions were restored and the column re-equilibrated for 4 minutes. The MS method consisted of a full scan in negative polarity mode at 70,000 resolution setting (at 200 m/z) with the mass range set between 380 and 2,400 m/z and AGC (acquisition gain control) target set at 3 x 106 and in-source CID set at 20 eV. Maximum injection time was set at 50 ms using 1 Microscan. Tune settings were as follows: spray voltage was 3.5 kV, sheath gas and auxiliary gas flows were set at 40 and 10 arbitrary units, respectively, the capillary temperature was 320 °C and probe heater temperature was 400 °C while S-lens RF voltage was 50 V.

1.4.4 *Statistical analysis*

Data acquisition, processing, and reporting collection were provided using Xcalibur ™ (Thermo Fisher Scientific, MA, USA). Progenesis QT (Waters TM) - a next-generation LC-MS multivariate data analysis software enabled glycans abundance normalization and relative quantification across groups. Putative *N*-glycans monosaccharide compositions were manually determined using GlycoWorkbench 3.0 tool based on given m/z values and considering mass accuracies below 20 ppm. The annotated fingerprints (experimental masses obtained from the MS fragmentation spectra) were searched against the database, deriving *N*-glycan structures (217). Interpretation of the relationship between the experimental groups was estimated by Principal Component Analysis (PCA), allowing the visualization of multivariate information. Statistical analysis and graph design were conducted with GraphPad Prism (GraphPad Software, Inc., San Diego, CA). Differences between groups were calculated using the Kruskal-Wallis test, followed by a Dunn's Multiple Comparison Test to correct for multiple comparisons. Normalized relative abundance profiles are presented as means of three technical replicates per group. $P \leq 0.05$ were accepted as significant.

1.5 *Proteomic analysis*

1.5.1 *Sample preparation for OSCC spectral library*

To accomplish Objective 3, a salivary protein spectral library was created through the identification of proteins present in a pool of saliva from OSCC patients. 10 out of the 165 individuals who have been sampled twice were selected, including 3 early and 7 advanced OSCC cases. After one freeze-thaw cycle, sample contents were mixed gently with a Pasteur pipette and centrifuged at 1500 rpm at 4°C for 15 min to achieve a homogeneous suspension. To increase the sensitivity of the protein content assay, samples were diluted in a 1:20 ratio with MQ H_2O.

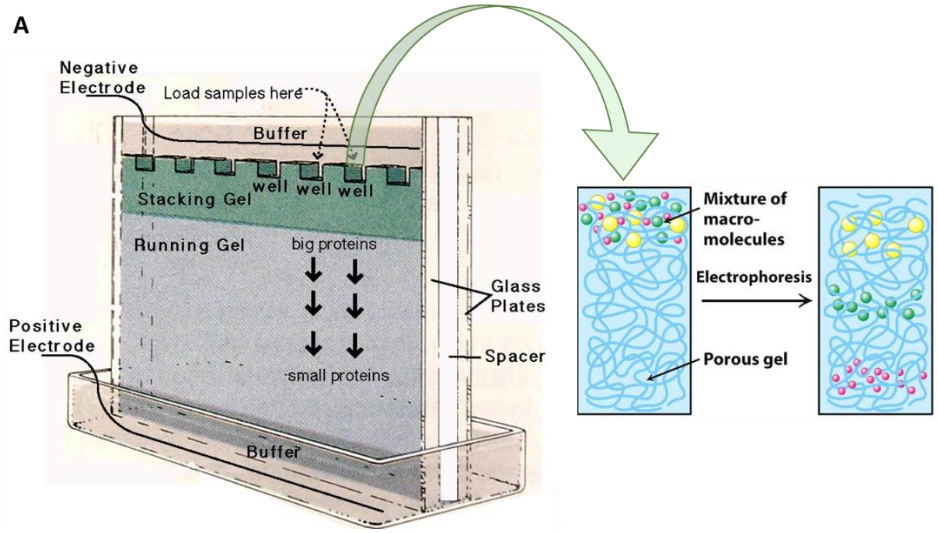
- Bradford assay

  Total protein concentrations of each one of the 10 samples were determined using the Bradford assay (215). It is a colorimetric method based on an absorbance shift of the Coomassie Brilliant Blue G-250 (Bio-Rad, US) protein dye, which forms a strong, noncovalent complex with the protein's carboxyl group by van der Waals force and amino group through electrostatic interactions. The binding of the protein stabilizes the blue form of the Coomassie dye and thus the amount of the complex present in the solution is a measure for the protein concentration, estimated by an absorbance reading. Perkin-Elmer VICTOR X3 ^{TM} spectrophotometer with an absorbance ratio 450/590 nm was implicated to evaluate the samples, standardized against 2mg/ml bovine gamma globulin (BGG), placed in a 96 well Greiner clear plate (218). 5 µg of protein from each sample were put together to form a total 50 µg protein pool. Before electrophoresis, the neat protein pool was treated with 5x Laemmli buffer containing β-mercaptoethanol (1:1) and kept in a heating block at 95°C for 5 min to reduce the intra and inter-molecular disulfide bonds. The whole amount was run into SDS PAGE.

    - SDS PAGE

  Polyacrylamide Gel Electrophoresis (PAGE) is a commonly employed technique for separating macromolecules, such as proteins. Electrophoresis is in general the process of applying an electric field to move charged molecules through a solution. In this approach, the mobility of a charged molecule is directly proportional to its net charge and the resistance of the solution through which it is moving. Sodium dodecyl sulfate (SDS) is an

amphipathic detergent. It has an anionic headgroup and a lipophilic tail. It binds non-covalently to proteins, with a stoichiometry of around one SDS molecule per two amino acids. SDS causes proteins to denature and disassociate from each other (excluding covalent cross-linking). Then proteins become negatively charged macromolecules with comparable hydrodynamic qualities, the mobility of which depends only on their molecular mass. In the presence of SDS, the intrinsic charge of a protein is masked. During SDS PAGE, as the current continues to flow, the proteins migrate into the resolving gel toward the anode (Fig 17A). Polyacrylamide gel with 1.5 mm thickness has been created in two layers, 12% "resolving" (“running”) gel on the bottom, and a narrow 4% "stacking" gel on top and kept between short and spacer glass plates forming the "gel cassette". Once polymerized, the gel cassette was mounted to an electrode assembly and placed into an electrophoresis tank (Fig.17B). Sample mix was loaded along with a marker for molecular weight determination (10-175kDa Pink Prestained Protein Marker; Tris-Glycine 15%, NipponGenetics^{TM}). The gel was run at 20mA constant voltage for 1h, removed from the cassette, and stained with Coomassie blue dye for 20 min. Excess dye removal was achieved by gel incubation in decolourant solution for approximately 2h.



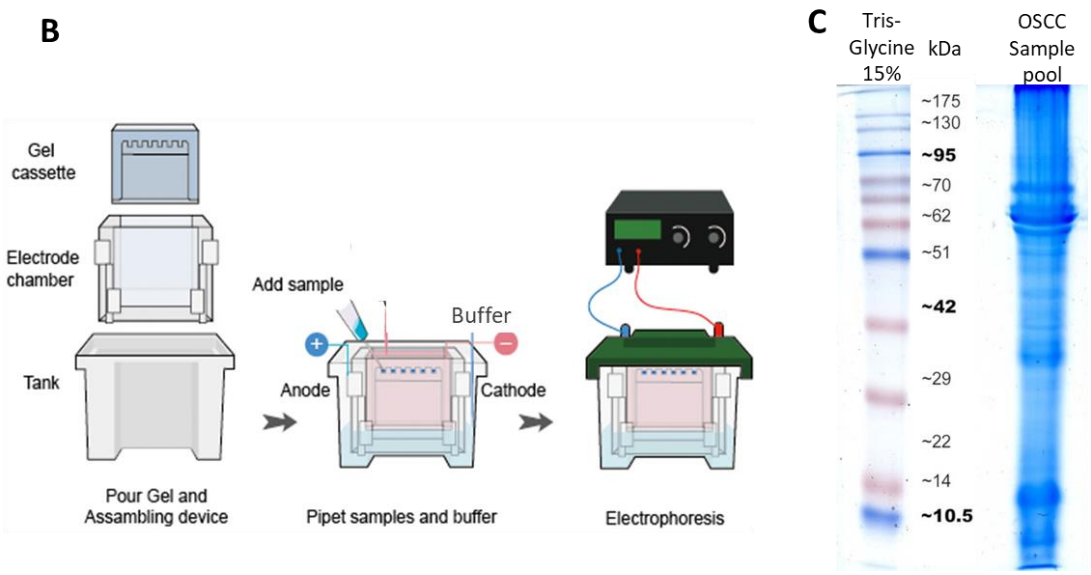


Figure 17. An illustration of an apparatus used for SDS PAGE. **(A)** When a sample is added into the stacking gel well, as the current flows, proteins separate during migration toward the positive electrode in a size-dependent manner. The smallest ones move faster through the running gel. **(B)** The gel cassette is mounted to an electrode chamber and placed into a tank, filled with electrophoresis buffer. The tank is closed with a lid, aligning appropriately the electrodes, and connected to a power supply. **(C)** Coomasie blue dye enabled protein visualization as separated bands along with the gel.

The gel stripe containing the colored protein bands was cut into five pieces as following: fragment I: 1-14kDa, fragment II: 14-30kD, fragment III: 30-50kDa, fragment IV: 50-95 kDa, and fragment V: 95-175 kDa proteins (Fig.17C). Samples were digested with sequencing grade trypsin (Promega) as described elsewhere (219). The digestion mixture was dried in a vacuum centrifuge, resuspended in 20 µL of 2% ACN, 0.1% TFA.

1.5.2 *Sample preparation for protein relative quantitation*

Protein differential expression was estimated in saliva samples obtained from 40 individuals, equally distributed into four groups according to the previously described inclusion criteria, as follows: 10 Controls, 10 PVL, 10 early (e), and 10 advanced (adv) OSCC stages. After one freeze-thaw cycle, sample contents were processed as per chapter 1.5.1 20 µg of total protein per sample were loaded in SDS PAGE and run without resolving within the conditions described above. Samples contained in a single gel band were digested with sequencing grade trypsin (Promega) as described elsewhere (219). The digestion mixture was dried in a vacuum centrifuge, resuspended in 20 µL of 2% ACN, 0.1% TFA. A pool of

samples from each of the four groups was evaluated as a quality control test, before the quantitation of the individual samples.

1.5.3 *LC-MS/MS technology*

Proteomics aims to completely identify and quantify the entire protein samples of interest. LC-MS is an analytical chemistry technique that combines the physical separation capabilities of liquid chromatography with the mass analysis capabilities of mass spectrometry (Fig.18). While LC separates mixtures with multiple components, MS provides structural identity and quantitative data of the individual components with high molecular specificity and detection sensitivity. MS measures the intact molecule, while tandem MS (MS/MS or MS2) refers to the analysis of the fragments (productions) that compose this molecule. This analysis requires the machine to isolate the entire particle and to break it, using electrical current and gas to produce specific signature ions from the parent molecule. These ions can then be used to identify and quantify a molecule. Data-dependent acquisition (DDA) is the mode of data collection where many peptides within a certain mass range are fragmented in tandem mass spectrometry. In DDA mode, the mass spectrometer selects the most intense peptide ions in the first stage of tandem mass spectrometry, and then they are fragmented and analysed in the second stage of tandem MS.

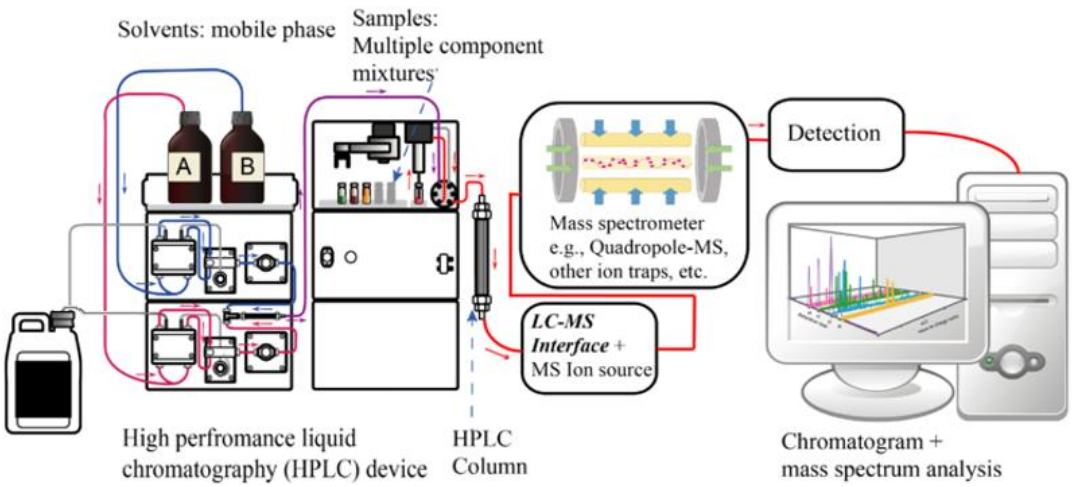


Figure 18. Functional diagram of a coupled LC-MS system. Adapted after Nagaraj *et al*.,2009 (183).

1.5.4 *Construction of salivary protein spectral library*

LC-MS were performed where 5 µl of gel sliced sample pool was loaded onto a trap column (NanoLC Column, 3µ C18-CL, 350 µmx 0.5mm; Eksigen) and desalted with 0.1% TFA at 3 µl/min during 5 min. The peptides were then loaded onto an analytical column (LC Column, 3 µ C18-CL, 75 umx12cm, Nikkyo) equilibrated in 5% acetonitrile 0.1% FA (formic acid). Elution was carried out with a linear gradient of 5a40% B in A for 120 min. (A: 0.1% FA; B: ACN, 0.1% FA) at a flow rate of 300nl/min. Peptides were analysed in a mass spectrometer nanoESI qQTOF (5600 TripleTOF, SCIEX). The sample was ionized applying 2.8 kV to the spray emitter. Analysis was carried out in a DDA mode. Survey MS1 scans were acquired from 350–1250 m/z for 250 ms. The quadrupole resolution was set to 'UNIT' for MS/MS experiments, which were acquired 100–1500 m/z for 50 ms in 'high sensitivity mode. Following switch criteria were used: charge: 2+ to 5+; minimum intensity; 70 counts per second (cps). Up to 50 ions were selected for fragmentation after each survey scan. Dynamic exclusion was set to 15 s. The system sensitivity was controlled with 2 fmol of 6 proteins (LC Packings). ProteinPilot^{TM} (v 5.0) search engine AB-SCIEX) was used for protein identification analysis.

1.5.5 *SWATH - MS quantification*

In the data-independent acquisition (DIA) mode, for each cycle, the instrument focuses on a narrow mass window of precursors and acquires MS/MS data from all precursors detected within that window (Fig.19). This mass window is then stepped across the entire mass range, systematically collecting MS/MS data from every mass and all detected precursors. The most common method to generate DIA data is called Sequential Windowed Acquisition of All Theoretical Fragment Ions (SWATH) in which the mass spectrometer divides the mass range into small mass windows. The SWATH-MS data consists of highly multiplexed fragment ion maps that are recorded over the user-defined mass precursor mass range and chromatographic separation.

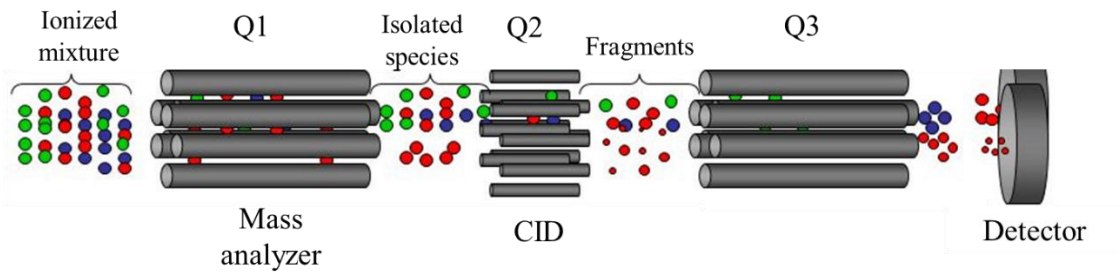


Figure 19. Triple quadrupole mass spectrometer. The ionized mixture of peptides enters the first quadrupole/mass analyser (Q1) in which the voltage settings are set to allow ions of a specific m/z value to pass through, selecting the ions of interest (MS1). The precursor ions then collide with argon gas in Q2 for fragmentation by Collision Induced Dissociation (CID). The third quad (Q3) scans repeatedly over the mass range to detect the fragment ions (MS2). Under defined kinetic energy, the flight time of the particle before it reaches the detector (at a set distance) directly relates to its mass-to-charge ratio (m/z). The separated ions are detected and this signal is sent to a data system where m/z ratios are stored together with their relative abundance for presentation in the format of an m/z spectrum. Modified after Ö Christophe D. Masselon, CEA Grenoble.

- For quality control test, 5 µL of sample pool per group was loaded onto a trap column (NanoLC Column, 3µ C18-CL, 350 µm x 0.5 mm; Eksigen) and desalted with 0.1% TFA at 3 µl/min during 5 min. The peptides were then loaded onto an analytical column (LC Column, 3 µ C18-CL, 75 umx12cm, Nikkyo) equilibrated in 5% acetonitrile 0.1% FA (formic acid). Elution was carried out with a linear gradient of 5-40% B in A for 120 min. (A: 0.1% FA; B: ACN, 0.1% FA) at a flow rate of 300nl/min. Peptides were analysed in a mass spectrometer nano Electrospray ionization (ESI) quadrupole time of flight (qQTOF) (5600 TripleTOF, SCIEX). The sample was ionized applying 2.8 kV to the spray emitter. DIA mode has been activated for analysis. Survey MS1 scans were acquired from 350–1250 m/z for 250 ms. The quadrupole resolution was set to 'UNIT' for MS/MS experiments, which were acquired 100–1500 m/z for 50 ms in 'high sensitivity mode. Following switch criteria were used: charge: 2+ to 5+; minimum intensity; 70 counts per second (cps). Up to 50 ions were selected for fragmentation after each survey scan. Dynamic exclusion was set to 15 s. The system sensitivity was controlled with 2 fmol of 6 proteins (LC Packings).

- For individual sample protein quantification, 5 µL of each sample was loaded onto a trap column (LC Column, 12 nm, 3 µ Triart-C18, 0.5 x5.0 mm; YMC) and desalted with 0.1% TFA at 10 µL/min during 5 min. The peptides were then loaded onto an analytical column (LC Column, Luna Omega 3 µm Polar C18, 150 x 0.3 mm, Capillary Phenomenex) equilibrated in 3% acetonitrile 0.1% FA (formic acid). Elution was carried out with a linear gradient of 3a353-35% B in A for 45 min (60 min total). (A: 0.1% FA; B: ACN, 0.1% FA) at a flow rate of 300nl/min. Peptides were analysed in a mass spectrometer microESI qQTOF (6600plus TripleTOF, ABSCIEX). The sample was ionized in a Source Type: Optiflow 1-50uL Micro applying 4.5 kV to the spray emitter. The analysis was performed in a DIA mode. Survey MS1 scans were acquired from 400–1250 m/z for 250 ms. 100 variable windows from 400 to 1250 m/z were acquired throughout the experiment. The total cycle time was 2.79 secs. The quadrupole resolution was set to 'UNIT' for MS2 experiments, which were acquired 100–1500 m/z for 25 ms in high sensitivity mode. The samples were acquired in a random order to avoid bias in the analysis.

1.5.6 *Data analysis*

ProteinPilot default parameters were used to generate a peak list directly from 5600 TripleTof wiff files. The Paragon algorithm (220) of ProteinPilot was used to search the Swissprot database (version 03-2018) with the following parameters: trypsin specificity, cys-alkylation, taxonomy restricted to humans, and the search effort set to through. The protein grouping was done by the Pro group algorithm. A protein group in a Pro Group Report is a set of proteins that share some physical evidence. Unlike sequence alignment analyses where full-length theoretical sequences are compared, the formation of protein groups in Pro Group is guided entirely by observed peptides only. Since the observed peptides are determined from experimentally acquired spectra, the grouping can be considered to be guided by the usage of spectra. Then, unobserved regions of protein sequence play no role in explaining the data. The wiff files obtained from the SWATH quantitation were analysed by Peak View 2.1 with the protein spectral library according to the scheme in Fig.20

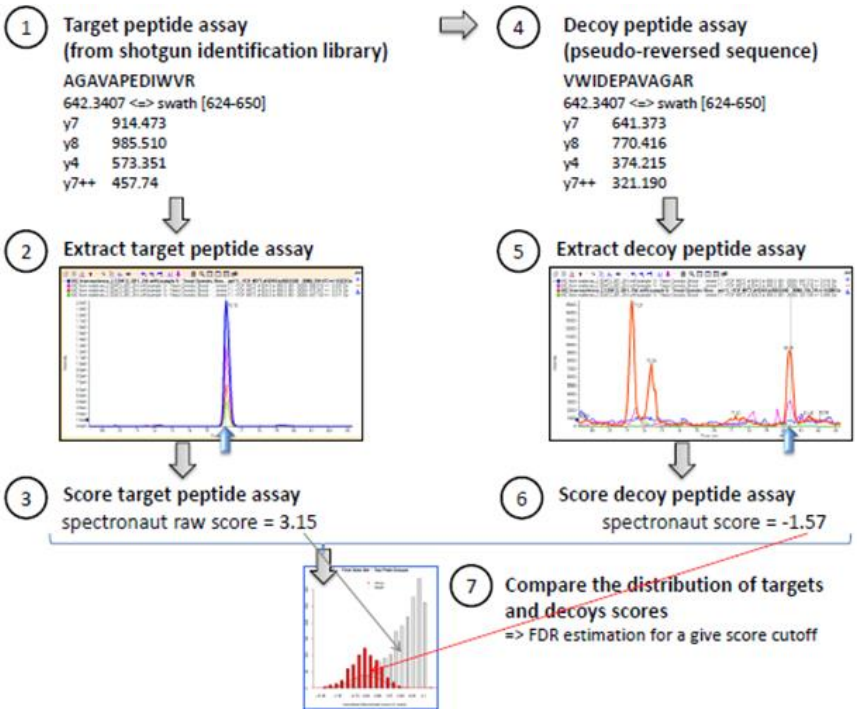


Figure 20. Peak View 2.1 analysis of SWATH generated data

The retention times were aligned among the different samples using main protein peptides. After retention time calibration, peptide selection was performed according to the processing settings shown in Fig. 21.

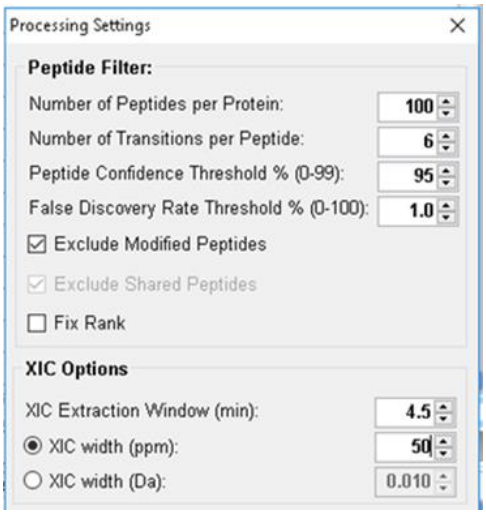


Figure 21. Processing setting of Peak View 2.1 for peptide selection

The cycle time used in the MS-MS/MS acquisition allows quantitating each peptide area with more than 7 points (Fig22).

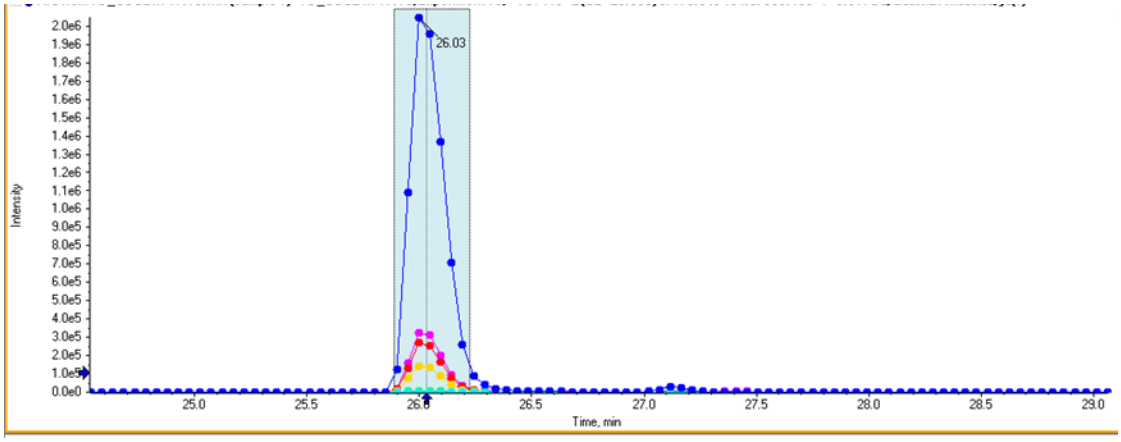


Figure 22. The area data obtained with Peak View is analyzed with Marker View (Sciex).

First, the protein areas calculated are normalized by the total sum of the areas of all the quantified proteins. Next, statistical tests of reduction of the dimensionality, principal component analysis (PCA), and discriminant analysis (DA) (both with Pareto scaling) were carried out. Due to the different variability of the groups, Welch T-tests were done for comparison. PCA aims to summarize the overall variability among individuals, which includes both the divergence between groups (i.e., structured genetic variability) and the variation occurring within groups ('random' genetic variability). To assess the relationships between different clusters, an adequate method should focus on between-group variability, while neglecting within-group variation. This is precisely the rationale of using Discriminant Analysis (DA) (221). The method, therefore, achieves the best discrimination of individuals into pre-defined groups.

1.5.7 *Biostatistics*

1.5.7.1 *Penalized linear regression methods*

A common challenge in biostatistical studies is to analyse in an efficient manner database in which the number of the variables of interest (predictors) is much greater than the number of the observations available. To overcome the limitation of traditional variable selection methods when the number of predictors is large, the methods for penalized regression and reduction of dimension (classification) have been developed. Cluster analysis was carried out with R 4.0.2 using *glmnet* package for LASSO (Least Absolute Shrinkage and Selection Operator) and Elastic Net Regularized Generalized Linear Models (222,223). The student's *t*-test was used to determine the statistical significance of differentially expressed proteins based on the average base 2 logarithmically transformed

(log2Med) relative abundance in pair-wise group comparisons. FRD adjusted *P*-value < 0.05 was accepted as statistically significant. Receiver operating characteristic (ROC) analyses were performed to evaluate the diagnostic value of discovered candidate biomarkers for OSCC, using MS Excel 2016. The area under the curve (AUC) of each ROC curve was obtained by numerical integration where significance was considered for a value great than 0.5

1.5.7.2 *Dimension reduction methodology*

Dimension reduction is the process of decreasing the number of random variables under consideration by obtaining a set of principal variables. Every dimension reduction method works in two steps: 1) obtaining the transformation of the predictor variables, and 2) estimation of the new regression model using these new predictors. The choice of these new predictors is usually main components or partial least squares (PLS). PLS combined with Discriminant Analysis (DA) is a method used when the response variable is qualitative. The normalized relative abundance values of all quantified proteins in control and pathology cases were transformed to Z scores and analysed via PLS-DA using the mixOmics package of R (223). Logarithmically transformed quantitative SWATH data was utilized to determine which proteins are the ones that best differentiate between early and advanced OSCC, PLV, and controls groups. The relationship between samples and proteins is represented graphically, using heat maps as an efficient method of visualizing intricate data sets organized as matrices (224). In a biological context, such a matrix is created by arranging the data in a way that each column contains the information from a single sample and each row corresponds to a single feature (protein). Heat maps permit to find quantitative patterns across proteins and biological samples simultaneously. Both samples and proteins are represented in order according to the result of the hierarchical classification.

1.5.7.3 *Functional analysis*

The STRING online software version 11.0 (225,226) was used to search for interaction relationships of the proteins differentially expressed in the OSCC compared to PVL and control groups, applying default settings and medium stringency. Biological process classifications were performed with the tools of the String database and enriched gene ontology (GO) terms provided by the software were also examined.

IV. *RESULTS*

1. *Analysis of salivary cytokines*

This analysis includes estimation and comparison of salivary levels of eight cytokines (IL-1α, IL-6, IL-8, IP-10, MCP-1, TNF-α, HCC-1, and PF-4) among patients with premalignant lesions (HL and PVL), early and advanced OSCC stages, and age-matched healthy controls.

1.1 *Participants' characteristics*

Summary of demographic and clinical characteristics of control individuals, leukoplakia, and OSCC patients are shown in Table 3, 4, and 5, respectively. The median age of the controls was 62 years (range: 50-82), of the HL patients was 68 years (range: 45-92), of the PVL was 67 years (range: 47-89), of the patients at early OSCC was 73 years (range: 39-95) and of those at advanced OSCC was 65 years (range: 39-95). There were no statistical differences among groups regarding age and sex. In the pathology cases, individuals' clinical characteristics such as oral lesion location, type, and histologic features were also registered.

Table 3. Characteristic features of the subjects comprising the control group.

	Control n=25
Age (median ± SD)	62 ±8.3
Sex	
Male	9 (36 %)
Female	16 (64%)
Tobacco smoking status	
Non-smokers	11 (44%)
Ex-smokers	10 (40%)
Current smokers	4 (16%)

All the patients comprising the homogeneous leukoplakia (HL) group exhibited a single lesion in the oral cavity presented as plain, white patches. Most of the patients included in the proliferative verrucous leukoplakia (PVL) cohort displayed multifocal rugged lesions, clinically presented as verrucous, mixed or erosive- ulcerative type (Table 4).

Table 4. Demographic and clinicopathological characteristics of homogeneous leukoplakia (HL) and proliferative verrucous leukoplakia (PVL) patients groups.

	HL (n=33)	PVL (n=33)
Age (median ± SD)	68 ± 12.1	67 ± 12.3
Sex		
Male	10 (30.3%)	13 (39.4%)
Female	23 (69.7%)	20 (60.6%)
Tobacco smoking status		
Non smokers	15 (45.5%)	24 (72.7%)
Current smokers	18 (54.5%)	9 (27.3%)
Associated location	n	n
Tongue	9	4
Floor of the mouth	2	2
Palate	3	1
Buccal mucosa	6	2
Gingiva	4	6
Multifocal	9	20
Lesion clinical form	n	n
Homogeneous white plaque	33 (100%)	-
Verrucous	-	14 (42.4%)
Erosive-ulcerative	-	5 (15.2%)
Mixed *	-	14 (42.4%)
Histologic features	n	n
Epithelial dysplasia	18 (54.5%)	14 (42.2%)
No epithelial dysplasia	15 (45.5%)	16 (48.5%)

*Homogeneous and verrucous regions

For the OSCC patients, information regarding tumor size, stage, and differentiation, as well as the presence of cervical adenopathy were recorded (Table 5).

Table 5. Demographic and clinic-pathological characteristics of patients at early and advanced OSCC stages.

	Early OSCC (n=33)	Advanced OSCC (n=33)
Age (median±SD)	73 ± 10.9	65 ± 15.6
Sex		
Male	13 (39.4%)	20 (60.6%)
Female	20 (60.6%)	13 (39.4%)
Tobacco smoking status		
Non smokers	23 (69.7%)	14 (42.4%)
Current smokers	10 (30.3%)	19 (57.6%)
Associated location	n	n
Tongue	17	7
Floor of the mouth	-	6
Palate	4	6
Buccal mucosa	3	5
Gingiva	8	8
Lip	1	1
Lesion clinical form		
Erythroplastic	6 (18.7%)	1 (3%)
Ulcerative	14 (43.7%)	25 (75.7%)
Exophytic	7 (21.8%)	1 (3%)
Mixed *	5 (15.6%)	6 (18.2%)
Lesion size	n	n
0-2 cm	12	6
2-4 cm	12	17
> 4 cm	-	6
Stage (TNM system)		
Tis; I; II	33	-
III; IV	-	33
Tumor differentiation		
well	26 (78.8%)	19 (57.6%)
not well **	7 (21.2%)	14 (42.4%)
Cervical adenopathy		
yes	6 (18.2%)	26 (78.8%)
no	27 (81.8%)	7 (21.2%)

*Ulcerative and exophytic; Tis – *carcinoma in situ*;

** include moderately and poorly differentiated tumors

1.2 *Variability analysis*

Variability, also called dispersion refers to how spread out a set of data is. Variability analysis were used to estimate how much cytokine expression levels differ from each other and allow further statistics to compare the investigated groups. Dispersion measures generated by R statistical programming of IL-1α, IL-6, IL-8, IP-10, MCP-1, TNF-α, HCC-1, and PF-4 expression in control and pathology groups are shown in a summary, including first and third quartiles, mean, median, minimum and maximum values, variance, standard deviation, as well as coefficient of variation in Table 6 (Supplementary Tables). Variation values (CV) of ≤ 1 suggest low variance data, pointing that cytokines present sufficient relative dispersion between themselves and between the groups, indicating heterogonous data within the groups.

1.3 *Principal component analysis*

The variability observed between samples and groups is largely determined by the nature of each group (principal component 1 - PC1) that represented 38.5% versus the intra-specific variability of each person (PC2) exhibited 16.9% (Fig.23). A clear variability increase is observed between the five groups, with the control group being the most homogeneous, which is expanding in HL and PVL samples, augments in the early OSCC, and even more so in the advanced OSCC group. This observation indicates that the pathologies present a clear dispersion with respect to the normal (control) pattern and between them as its severity increases. Some outliers are spotted, especially in the advanced OSCC group. However, including all the samples' values allows maintaining variability as a factor to be considered in the further analysis.

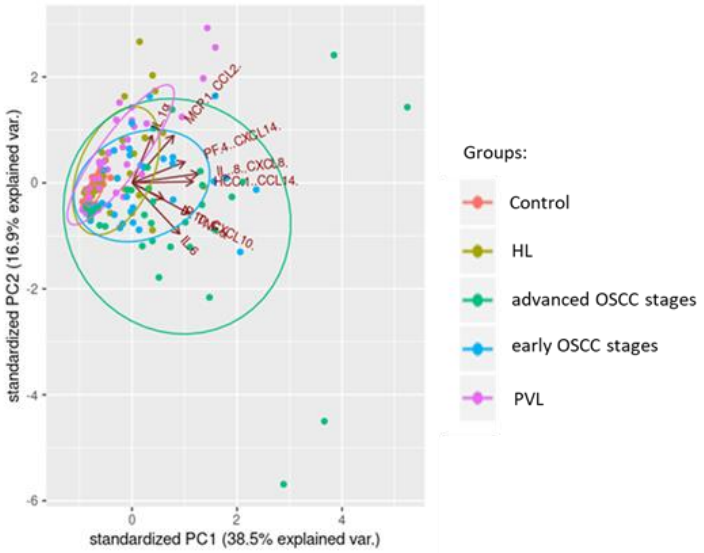


Figure 23. Principle component plot based on IL-1α, IL-6, IL-8, IP-10, MCP-1, TNF-α, HCC-1 y PF-4 levels of detection (LOD) estimated with LuminexTM xMAP in Control (n=25), HL (n=33), PVL (n=33), early OSCC (n=33) and advanced OSCC (n=33).

1.4 *Cluster analysis*

A cluster analysis was carried out and a heat map was generated to graphically visualize the changes in cytokine values within the control, HL, PVL, early and advanced OSCC samples. In the heat map, each column contains the data from a single sample and each row corresponds to a single feature (cytokine). A heat map reorders the rows and columns so that rows with similar profiles are closer to one another, causing these profiles to be more visible. Each entry in the data matrix is displayed as a colour, marking it possible to view the patterns graphically. To represent visually, how the quantified analytes are related to the groups, a heat map has been created where concentration values of the studied cytokines are clustered in rows and the sample groups in columns (Fig.24). The scale indicates cytokine concentration in ranges from low (red) to high (blue) values. Based on cluster analysis, the values can discriminate the control from the pathologies group. The columns were clustered using the complete linkage with the Euclidean distance measure method by R software for statistical computing. IL-6, TNF-α, and HCC-1 present low expression profiles (red range) that contrast with PF-4 presenting a medium degree of expression (white), while IL-1α, IL-8, IP-10, and MCP-1 demonstrate higher value profile(blue range). Some intragroup variability is observed in each cohort, however, differences are notable between control and pathology groups, especially in the OSCC, for all analytes.

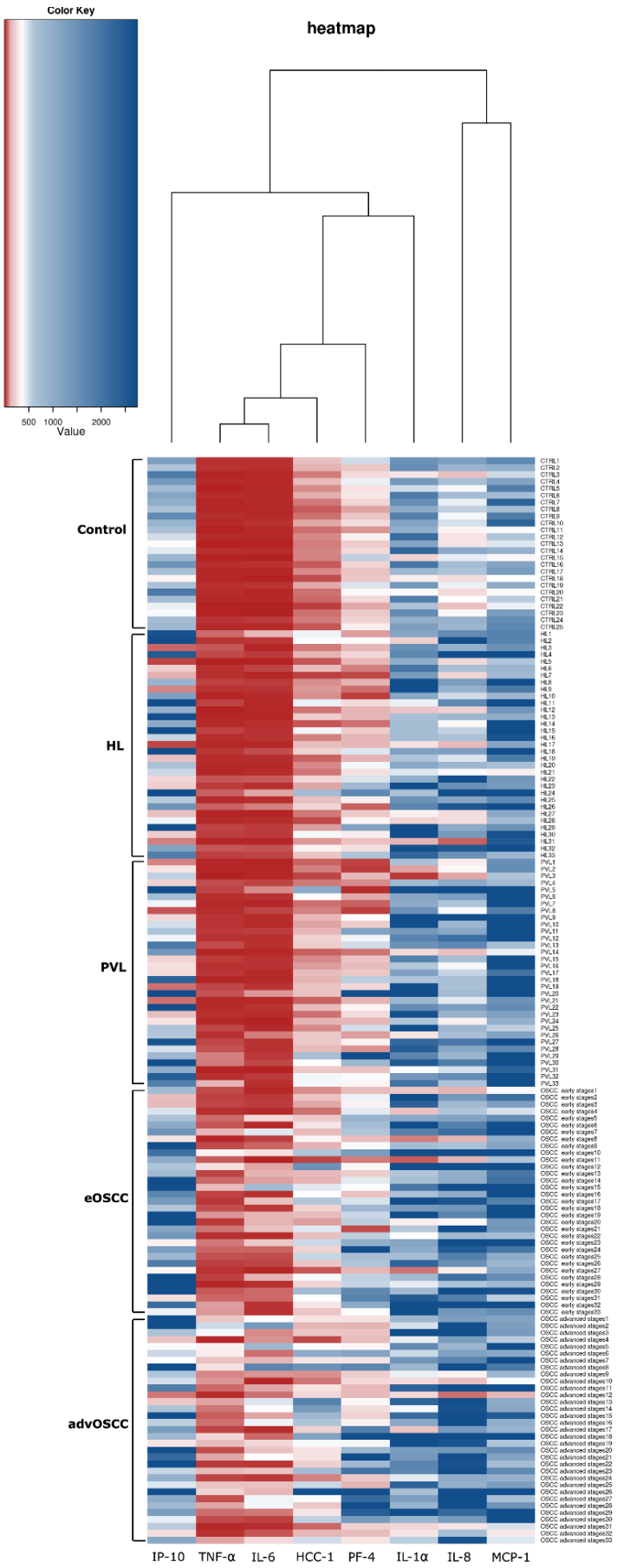


Figure 24. Cluster analysis and heat map of eight cytokines detected in saliva from patients with HL, PVL, early (e) and advanced (adv) OSCC stages, and age-matched controls. The colour key at the top of the heat map shows the Level of Detection (LOD) ranging from low (red) to high (blue) values

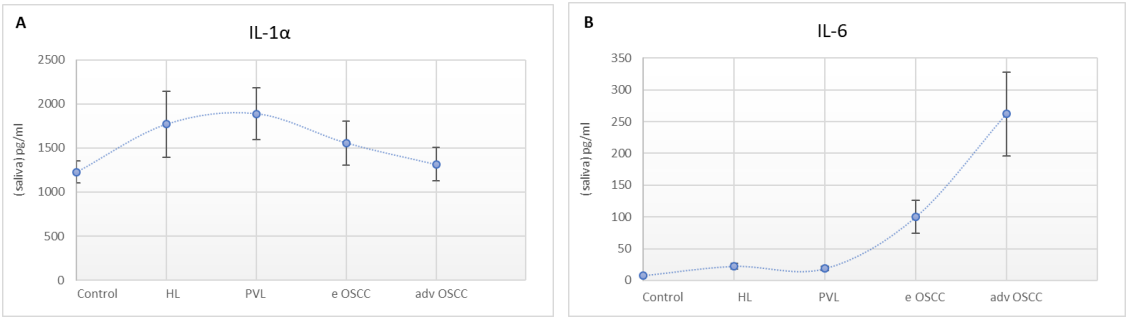
1.5 *Comparison of cytokine expression*

Salivary concentrations of IL-1α, IL-6, IL-8, IP-10, MCP-1, TNF-α, HCC-1, and PF-4 in patients with HL, PVL, at early and advanced OSCC stages and age-matched controls were estimated using the multiplexed bead-based assay. A summary of the mean levels (pg/mL) and SEM of each cytokine per group is presented in Table 7. The mean salivary concentrations of IL-6, IL-8, TNFα, HCC-1, and PF-4 were found to follow similar expression patterns of gradual increase from controls to leukoplakia (HL and/or PVL), to early and finally to advanced OSCC stages (Fig.25 B, C, F, G, and H, respectively). Transient alterations were observed in IL-1α, IP-10 (CXCL10), and MCP-1 (CCL2) concentrations (Fig.25 A, D, and E, respectively) within the groups.

Table 7. Salivary cytokines levels in pg/mL in controls, HL, PVL, early (e), and advanced (adv) OSCC patients; results are presented as the arithmetic mean and standard error of the mean (SEM)

	Control (n=25)	SEM	HL (n=33)	SEM	PVL (n=33)	SEM	e OSCC (n=33)	SEM	adv OSCC (n=33)	SEM
IL-1α	1227,24	122,90	1768,11	376,97	1886,48	294,87	1556,19	248,31	1313,79	189,58
IL-6	7,95	0,95	22,61	4,78	18,90	3,75	99,82	26,09	262,08	65,83
L-8	526,17	59,03	1382,92	279,60	1140,87	240,48	2567,01	549,72	4124,81	787,23
IP-10	884,97	93,39	1649,98	386,48	1157,80	251,30	1963,89	365,41	1556,53	332,42
MCP-1	1066,61	126,54	2535,35	372,84	3600,83	481,32	2560,73	370,26	2099,76	429,95
TNF-α	7,62	0,84	19,65	2,82	23,08	4,26	59,82	11,56	122,52	20,68
HCC-1	75,36	6,10	174,55	27,66	189,28	35,77	298,53	49,55	551,32	119,67
PF-4	253,74	23,98	293,92	60,53	454,78	146,61	642,00	104,85	1021,85	220,02

HL- homogeneous leukoplakia; **PVL**- proliferative verrucous leukoplakia; **e** – early; **adv**- advanced OSCC;



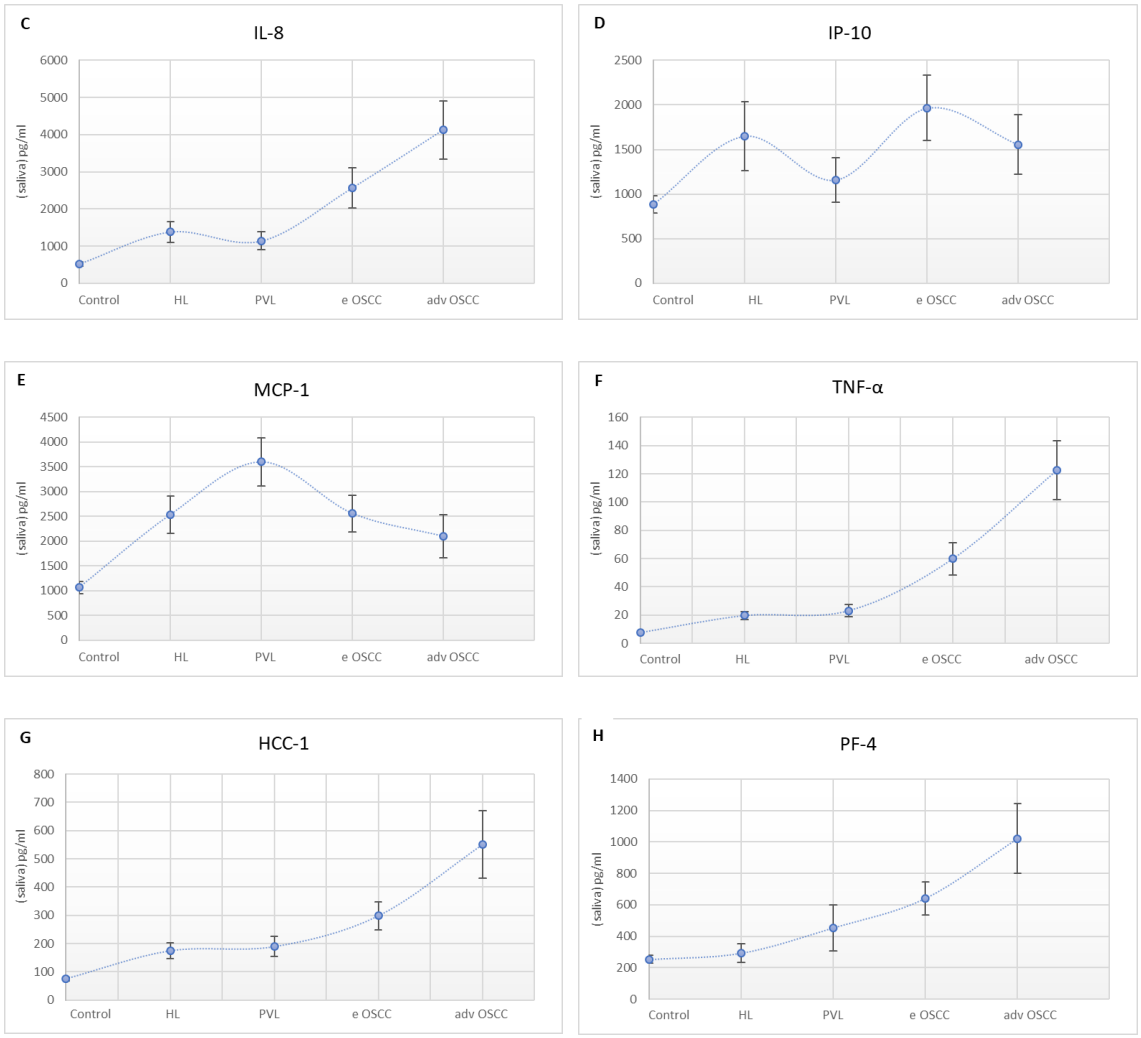
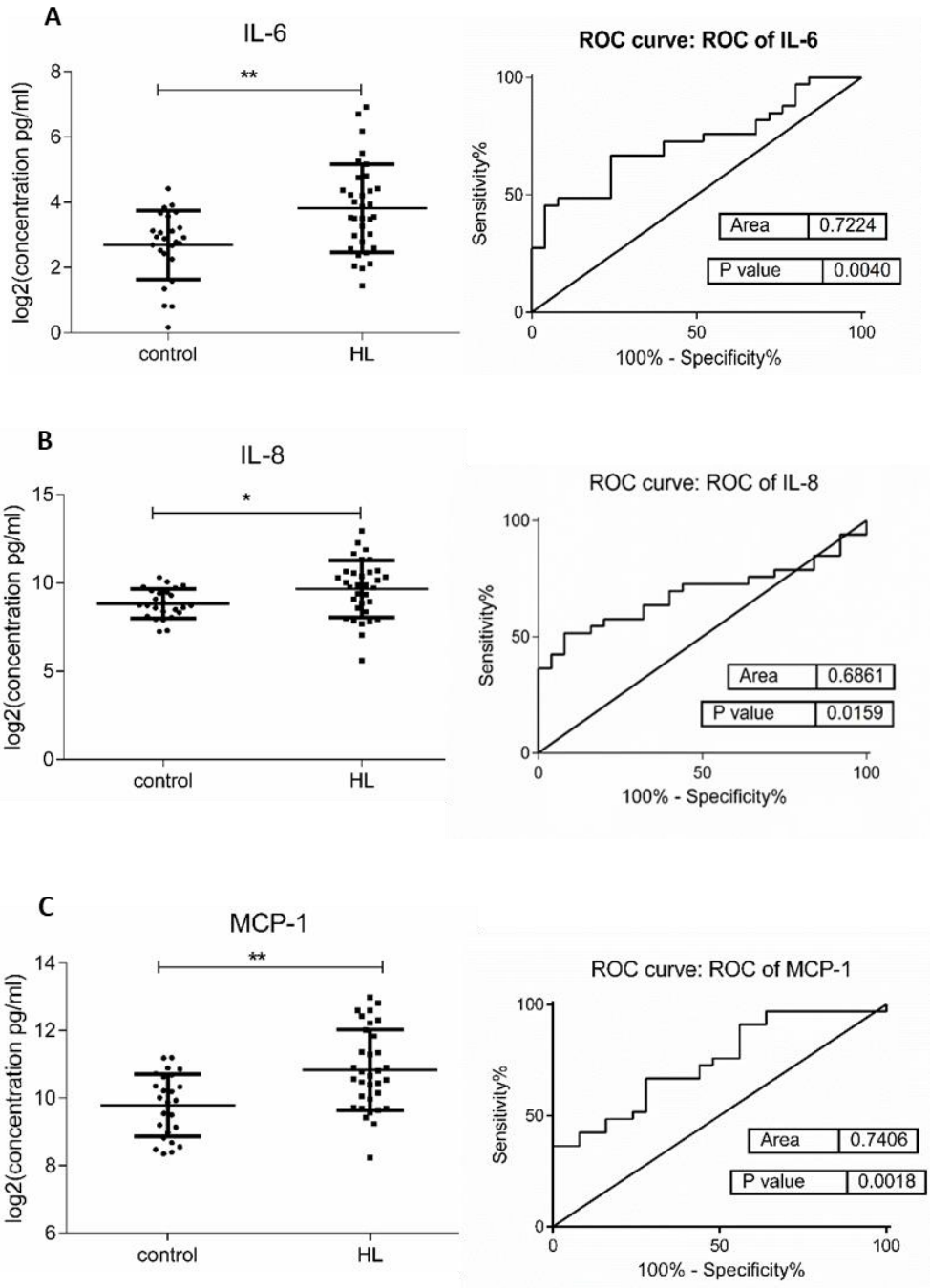


Figure 25. Salivary levels in pg/mL of **(A)** IL-1α, **(B)** IL-6, **(C)** IL-8, **(D)** IP-10, **(E)** MCP-1, **(F)** TNF-α, **(G)** HCC-1, and **(H)** PF-4. The blue lines indicate cytokine modulation trend from control to HL, to PVL, to early, and finally to advanced OSCC, where each dot represents the mean ±SEM of n=25 for control and n=33 for pathology groups and n is an average of two technical replicates.

Independent, nonparametric Mann-Whitney and Kruskal-Wallis tests followed by Dunn's multiple comparisons were utilized, when appropriate, to assess differences in logarithmically transformed (log2) cytokine mean levels between the groups. *P*-value ≤ 0.05 was considered for statistically significant and indicated with **** for $p \leq 0.0001$; *** for $p \leq 0.001$; ** for $p \leq 0.01$ and * for $p \leq 0.05$. Dot plots represent salivary levels of each analyte in log2 concentration (pg/mL) in all the comparisons. Prediction model for sensitivity and specificity performance of altered cytokines is depicted with ROC curves. Significance was considered for the area under the curve (AUC) value great than 0.5

1.5.1 *Control vs Leukoplakia*

The mean salivary concentrations of IL-6, IL-8, MCP-1, TNF-α, and HCC-1 marked significantly higher expressions in HL patients than in their control counterparts and AUC values of 0.72, 0.68, 0.74, 0.72, and 0.72, respectively (Fig.26 A, B, C, D, and E, respectively). No significant differences were found in IL-1α, IP-10, and PF-4 levels between the compared groups (Fig.26 F, G, and H, respectively).



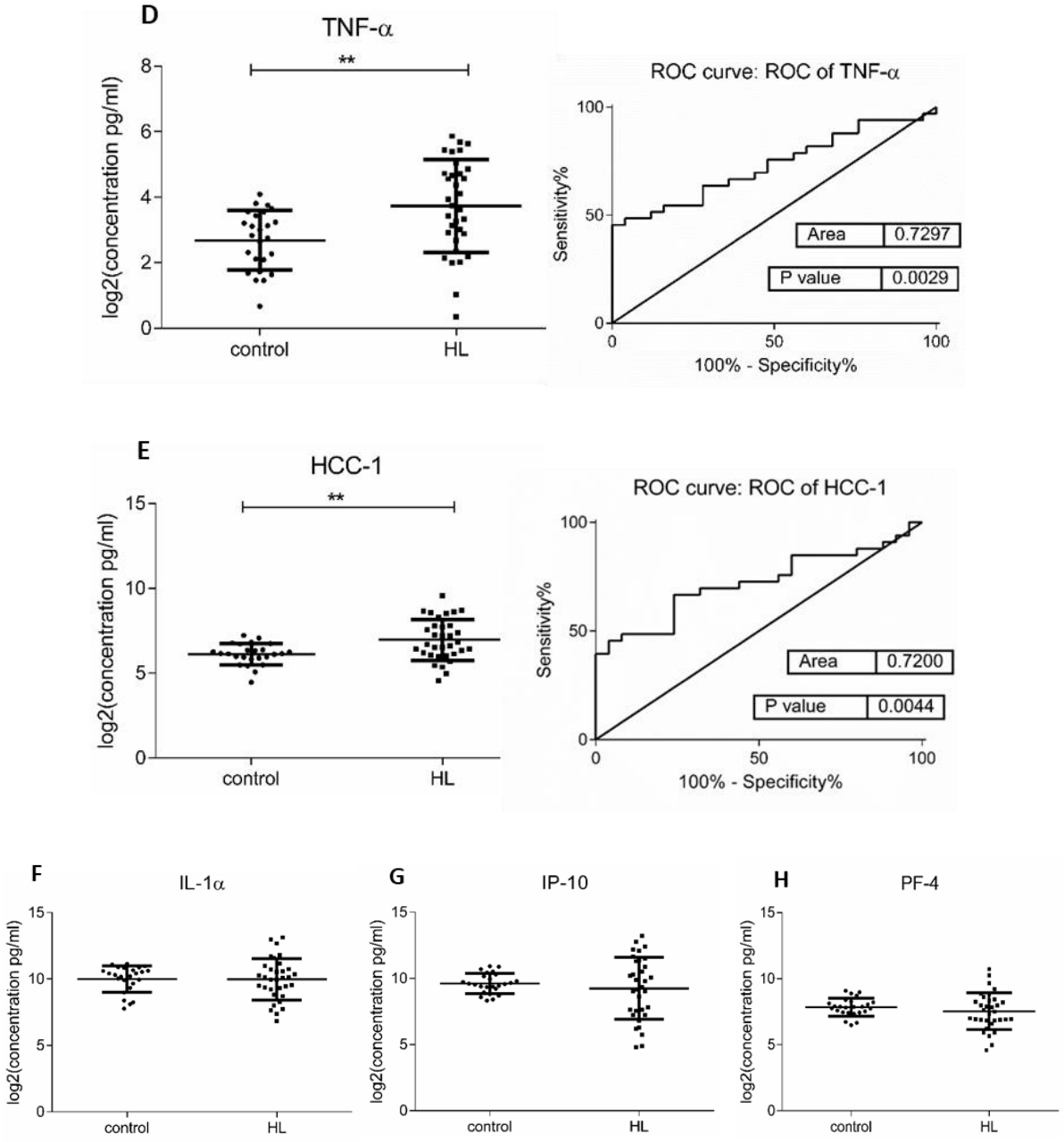
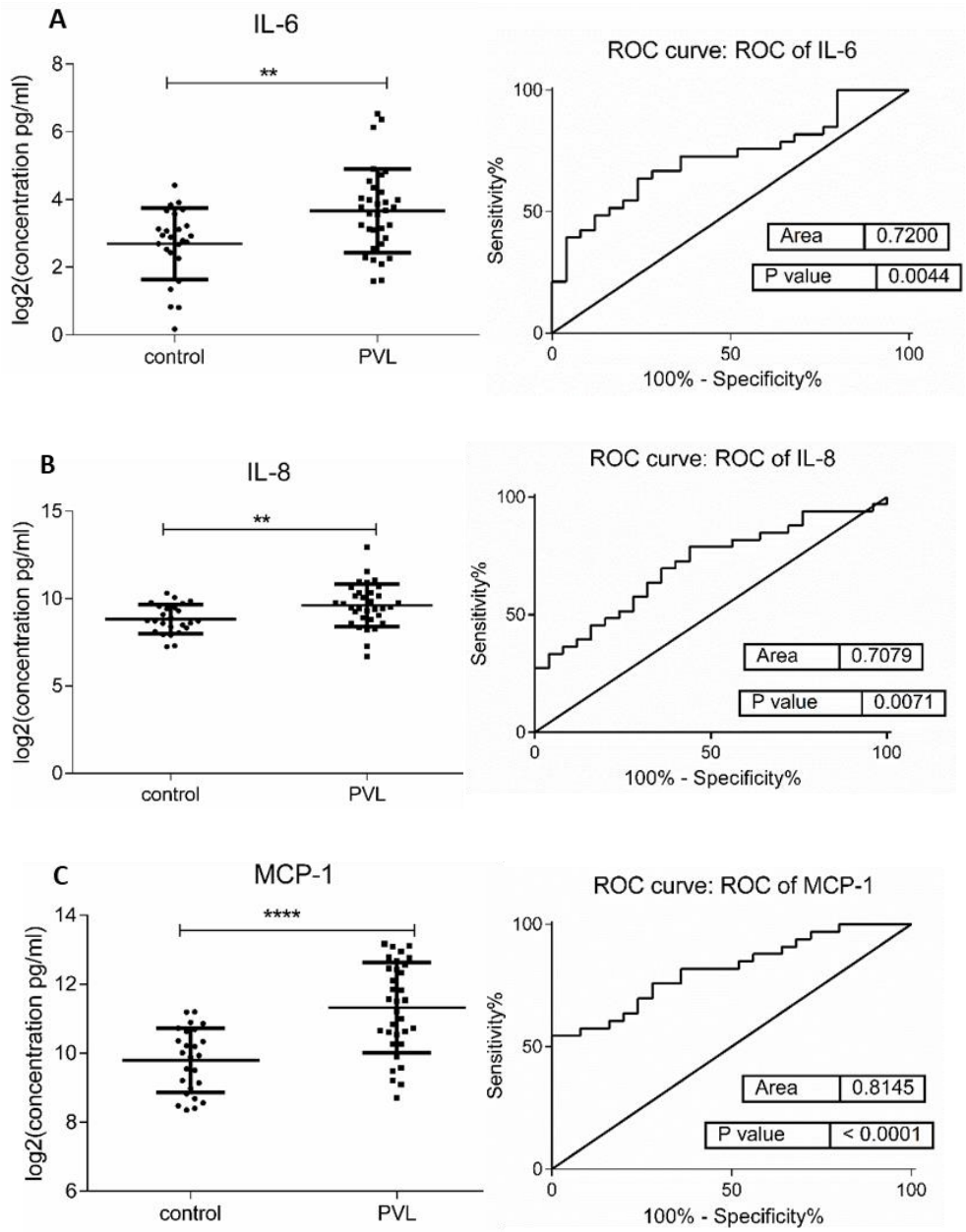


Figure 24. Comparison of cytokine levels in saliva of control volunteers and HL patients. Dot plot (left) of cytokine log2 levels (pg/mL) and ROC curves (right) of (**A**) IL-6, (**B**) IL-8, (**C**) MCP-1, (**D**) TNF-α and (**E**) HCC-1 (p = 0.01, 0.05, 0.01, 0.01 and 0.01, respectively). Dot plots of (**F**) IL-α, (**G**) IP-10 and (**H**) PF-4. Values represent mean ±SEM of n=25 (control group) and n=33 (HL group) where, n is an average of two technical replicates.

Similarly, in a collation of estimated cytokine expression between volunteers with PVL and healthy controls, salivary levels of IL-6, IL-8, MCP-1, TNF-α, and HCC-1 were seen significantly increased in favor of the patients with premalignant lesions (Fig.27 A, B, C, D, and E, respectively). The prediction model for sensitivity and specificity performance

revealed AUC values greater than 0.7. IL-1α, IP-10, and PF-4 showed no considerable changes among the compared cohorts (Fig.27 F, G, and H, respectively)



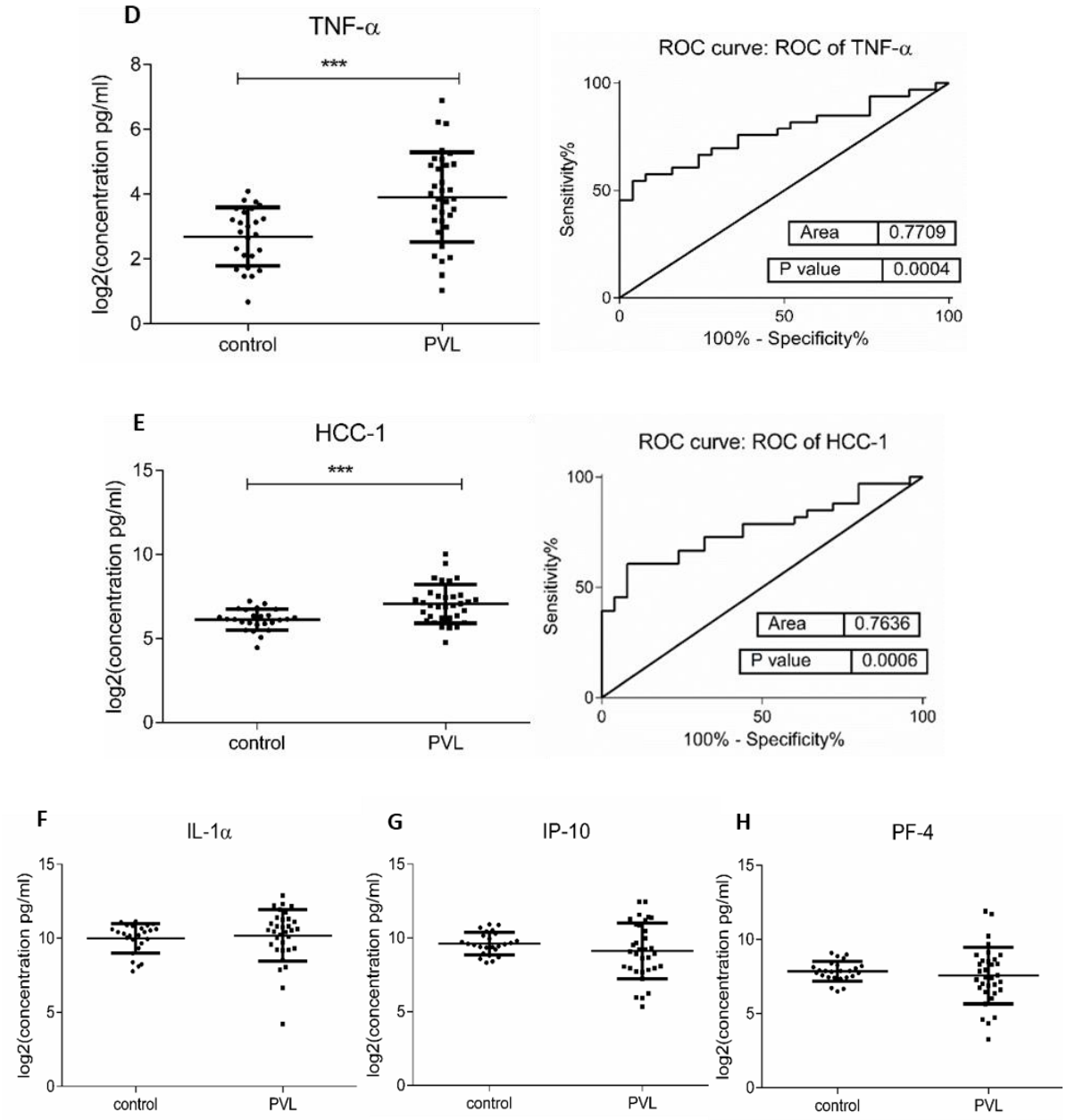
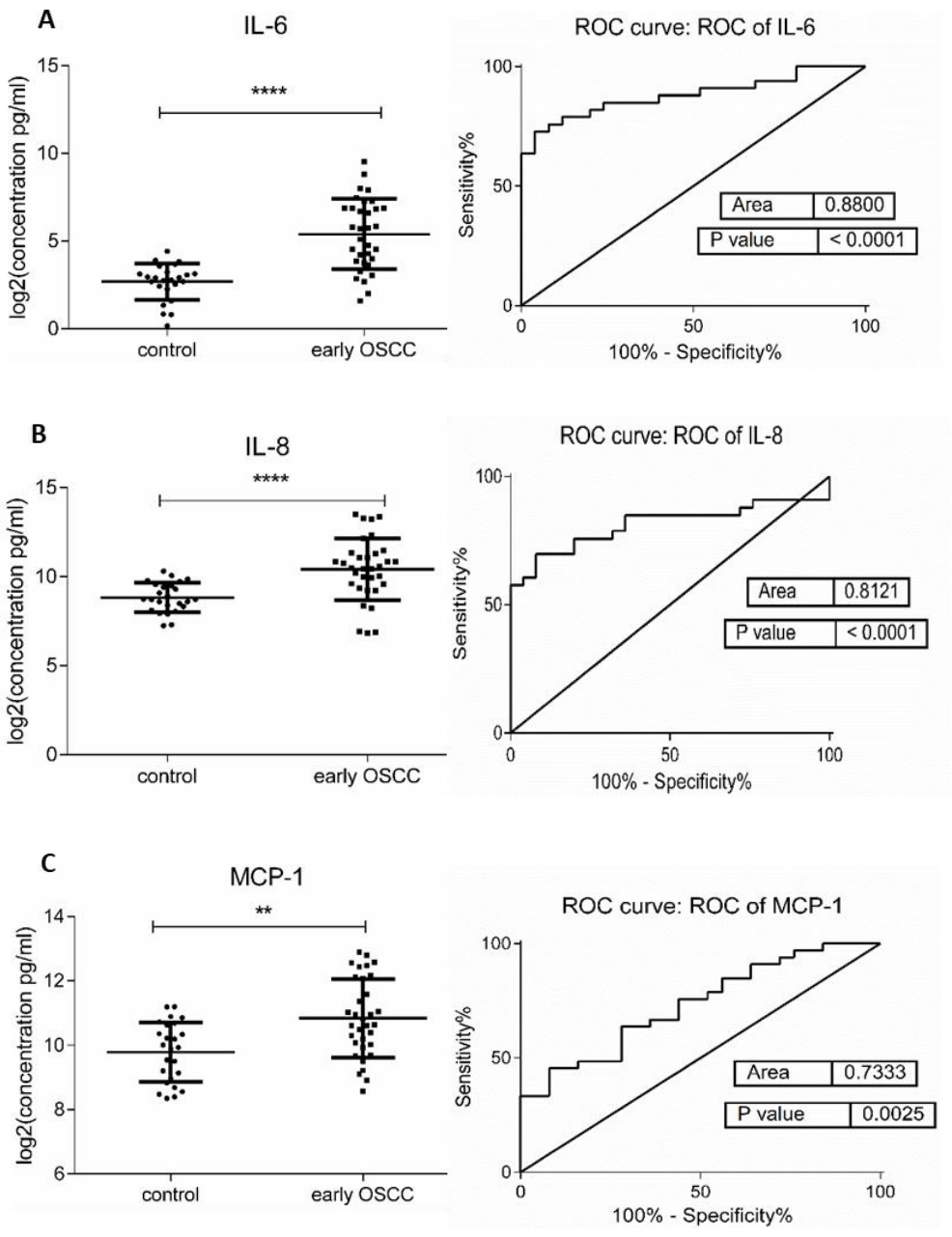


Figure 25. Comparison of cytokine concentrations in the saliva of and control volunteers and PVL patients. Dot plot (left) of cytokine log2 levels (pg/mL) and ROC curves (right) of (**A**) IL-6, (**B**) IL-8, (**C**) MCP-1, (**D**) TNF-α and (**E**) HCC-1 (p =0.01, 0.01, 0.001, 0.001 and 0.001, respectively). Dot plots of (**F**) IL-α, (**G**) IP-10, and (**H**) PF-4. Values represent mean ±SEM of n=25 (control group) and n=33 (PVL group), where n is an average of two technical replicates.

No significant differences were estimated in levels of the target cytokines among patients with HL and PVL. Dot plot graphs, representing the expression profiles of the 8 cytokines in the two clinical types of oral leukoplakia are shown in Figure 28 (Supplementary figures).

1.5.2 *Control vs OSCC*

A comparison of target cytokine concentrations revealed significant overexpression of IL-6, IL-8, MCP-1, TNF-α, HCC-1, and PF-4 and AUC values greater than 0.7 in patients at early OSCC compared to matched controls (Fig.29 A, B, C, D, E, and F, respectively). The above-mentioned cytokines, excluding MCP-1, exhibited a considerable increase also in advanced disease stages (Fig.30, Supplementary figures). IL-1α and IP-10 did not show notable alterations, independently of the OSCC clinical stage (Fig.29 G and H and Fig.30, Supplementary figures).



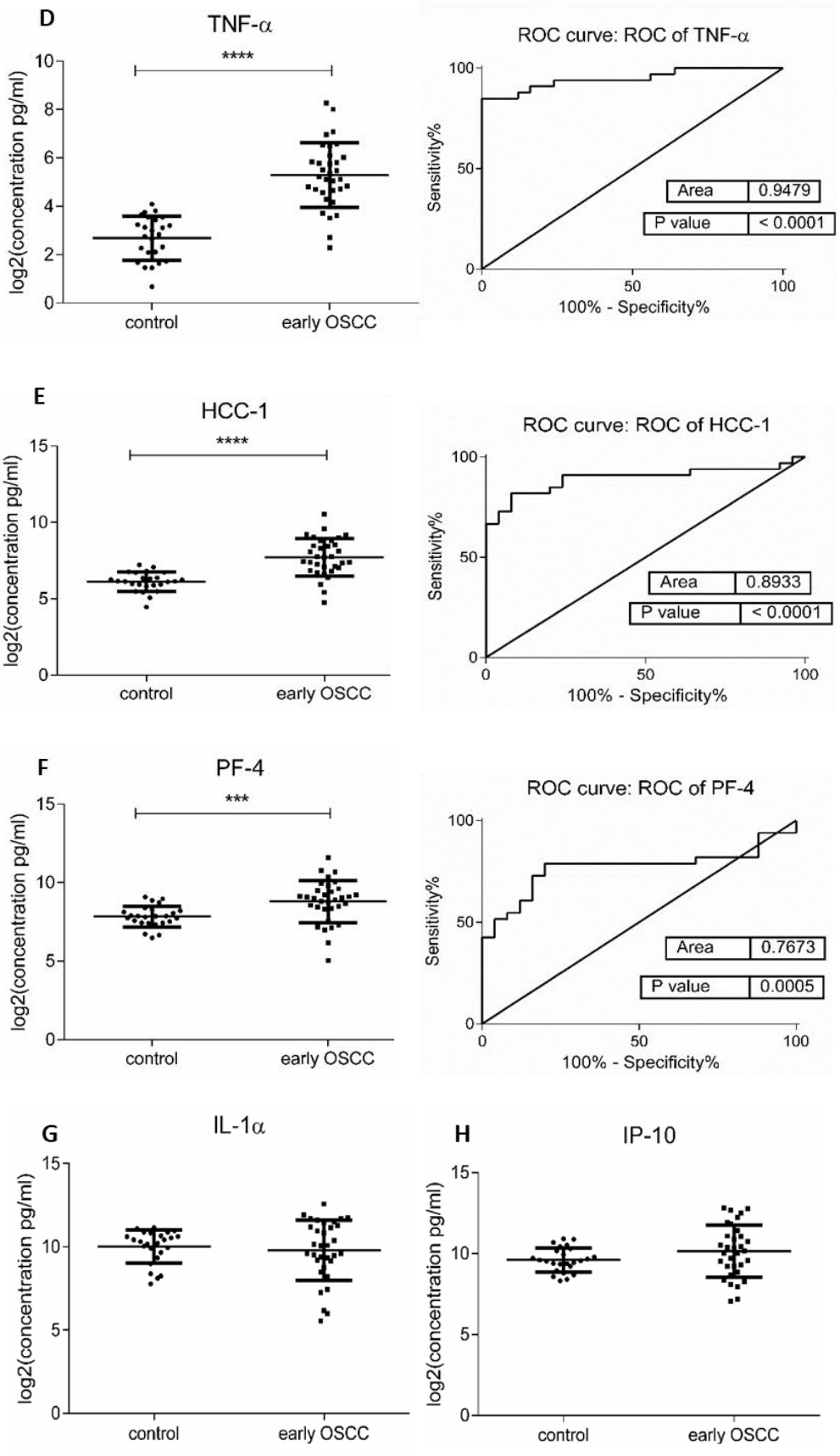
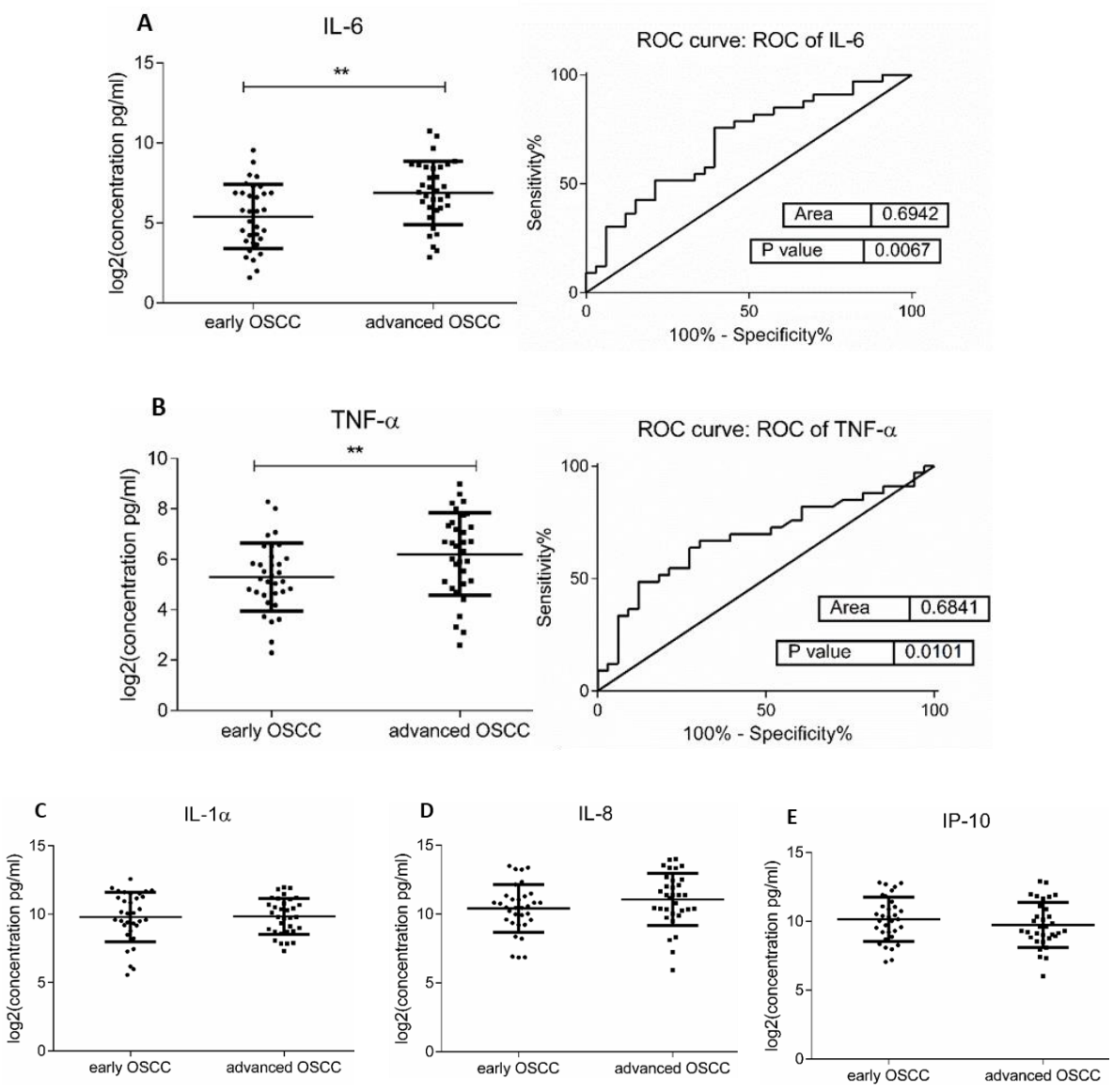

D
TNF-α
****
log2(concentration pg/ml)
control
early OSCC
ROC curve: ROC of TNF-α
Sensitivity%
100% - Specificity%
Area 0.9479
P value < 0.0001
E
HCC-1
****
log2(concentration pg/ml)
control
early OSCC
ROC curve: ROC of HCC-1
Sensitivity%
100% - Specificity%
Area 0.8933
P value < 0.0001
F
PF-4
***
log2(concentration pg/ml)
control
early OSCC
ROC curve: ROC of PF-4
Sensitivity%
100% - Specificity%
Area 0.7673
P value 0.0005
G
IL-1α
log2(concentration pg/ml)
control
early OSCC
H
IP-10
log2(concentration pg/ml)
control
early OSCC

Figure 29. Salivary cytokine expressions in patients at early OSCC stages and controls. Dot plot (left) of cytokine log2 levels (pg/mL) and ROC curves (right) of (**A**) IL-6, (**B**) IL-8, (**C**) MCP-1, (**D**) TNF-α and (**E**) HCC-1 and (**G**) PF-4 ($p \leq 0.0001, 0.0001, 0.01, 0.0001, 0.0001$ and 0.05, respectively). Dot plots of (**G**) IL-α and (**H**) IP-10. Values represent mean ±SEM of n=25 (control group) and n=33 (early OSCC group), where n is an average of two technical replicates.

1.5.3 *Early OSCC vs advanced OSCC*

Estimated levels of the 8 cytokines in the saliva of early OSCC were set against advanced OSCC. Comparisons of the analysed proteins between the aforementioned groups are presented in Figure 31. Considerably higher IL-6 and TNF-α with AUCs of 0.69 and 0.68, respectively were estimated in advance compared to early OSCC stages. IL-1α, IL-8, IP-10, MCP-1, HCC-1, and PF-4 concentrations did not differ significantly amongst patients at different OSCC clinical stages.



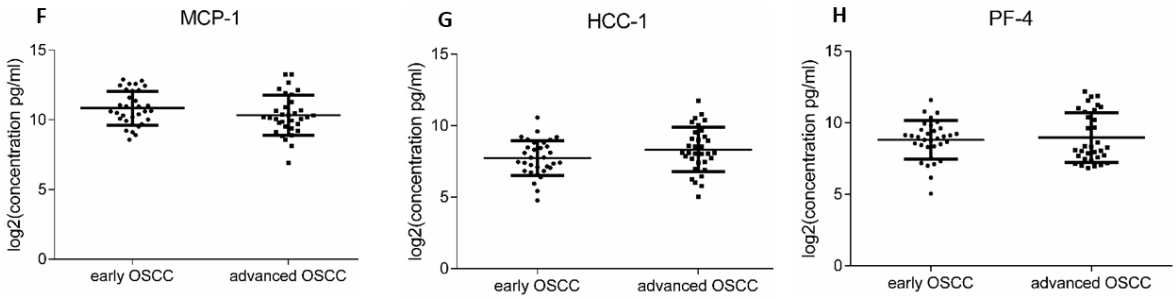
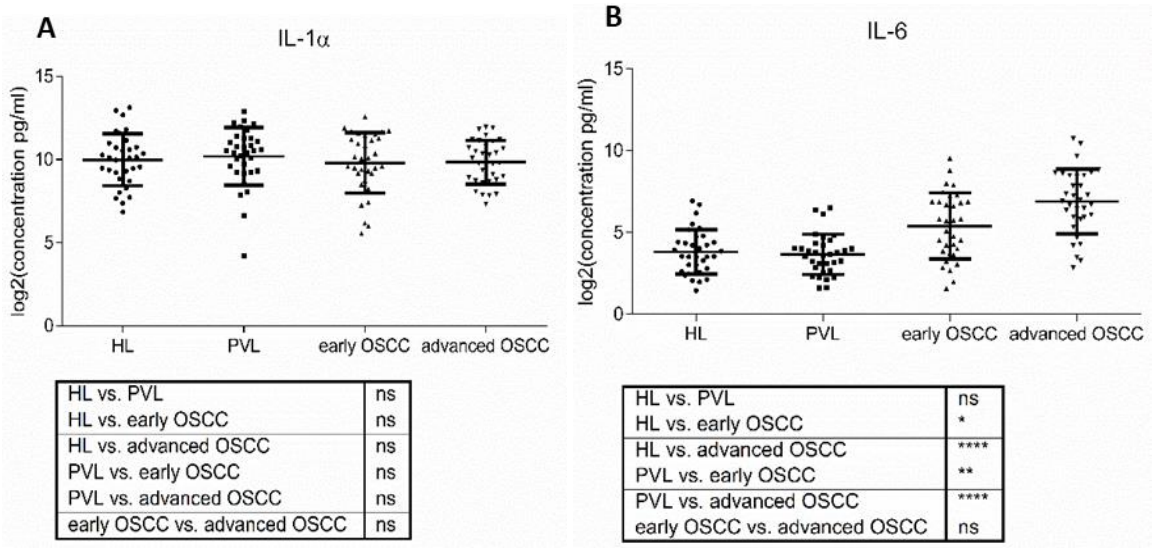


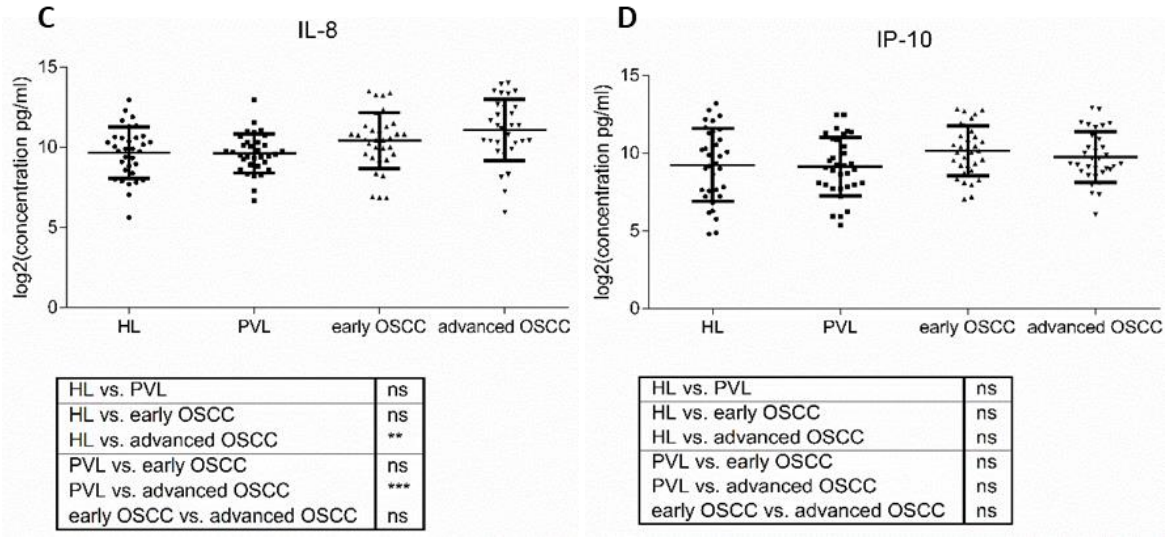
Figure 31. Comparison of salivary cytokine concentrations in patients at early and advanced OSCC stages. Dot plot (left) and ROC curve (right) showing significantly increased log2 levels (pg/mL) of (**A**) IL-6 and, (**B)** TNF-α (p = 0.01 and 0.01, respectively). Dot plots of (**C**) IL-1α, (**D**) IL-8, (**E**) IP-10, (**F**) MCP-1, (**G**) HCC-1 and (**H**) PF-4. Values represent mean ±SEM of n=33 (early OSCC) and n=33 (advanced OSCC), where n is an average of two technical replicates

1.5.4 *Leukoplakia vs OSCC*

Salivary profiles of the eight cytokines were compared between subjects with HL and PVL and at early and advanced OSCC stages. Results showed notably altered IL-6, IL-8, TNF-α, HCC-1, and PF-4 levels in pair-wise and multiple comparisons. Significantly increased concentrations of the aforementioned inflammatory agents were estimated in early (Fig.32, Supplementary figures) and advanced OSCC, compared to leukoplakia forms (Fig.33, Supplementary figures) collated to oral leukoplakia patients (HL&PVL). ROC analysis demonstrated significant AUC values greater than 0.7 for IL-6, TNF-α, and PF-4 at early cancer stages, while in progressed disease higher than 0.8 for IL-6 and TNF-α, and ≥ 0.7 for IL-8 and HCC-1. IL-α, IP-10, and MCP-1 exhibited similar expression patterns in patients with premalignant and cancerous lesions (Fig. 32 and 33, Appendix 1). Multiple comparisons of cytokine levels among the pathology cohorts are presented with dot plot graphs in Figure 34. It can be seen that IL-6 marked considerable differences within the four groups with p = 0.0001. Dun's test based on pair-pair comparisons confirmed significance in HL vs early OSCC (p = 0.05), HL vs advanced OSCC (p = 0.0001), PVL vs early OSCC (p = 0.01), and PVL vs advanced OSCC (p = 0.0001) (Fig.34B). For IL-8, important changes were detected with p <0.0003 and by post hoc test of HL vs advanced OSCC (p = 0.01) and PVL vs advanced OSCC (p = 0.001) (Fig.34C). Salivary TNF-α was found notably amended between the four groups with p <0.0001, while pair-pair comparisons revealed distinction of HL vs early OSCC (p = 0.05), HL vs advanced OSCC (p = 0.0001), PVL vs early OSCC (p = 0.01) and PVL vs advanced OSCC (p= 0.0001) (Fig.34F). HCC-1 and PF-4 exhibited

differential expression within the collated groups with p = 0.0002 and p = 0.0001, respectively. HCC-1 exhibited an increase in advanced OSCC against HL and PVL with p = 0.01 (Fig. 34G). PF-4 showed significant elevation in the OSCC groups collated to the ones of premalignant disorders, namely HL vs early OSCC (p = 0.01), HL vs advanced OSCC (p = 0.01), PVL vs early OSCC (p = 0.01) and PVL vs advanced OSCC (p = 0.05) (Fig34H). No appreciable alteration was assessed in the concentrations of IL-1α (Fig34A), IP-10 (Fig34D), and MCP-1 (Fig34E) across the pathologies.





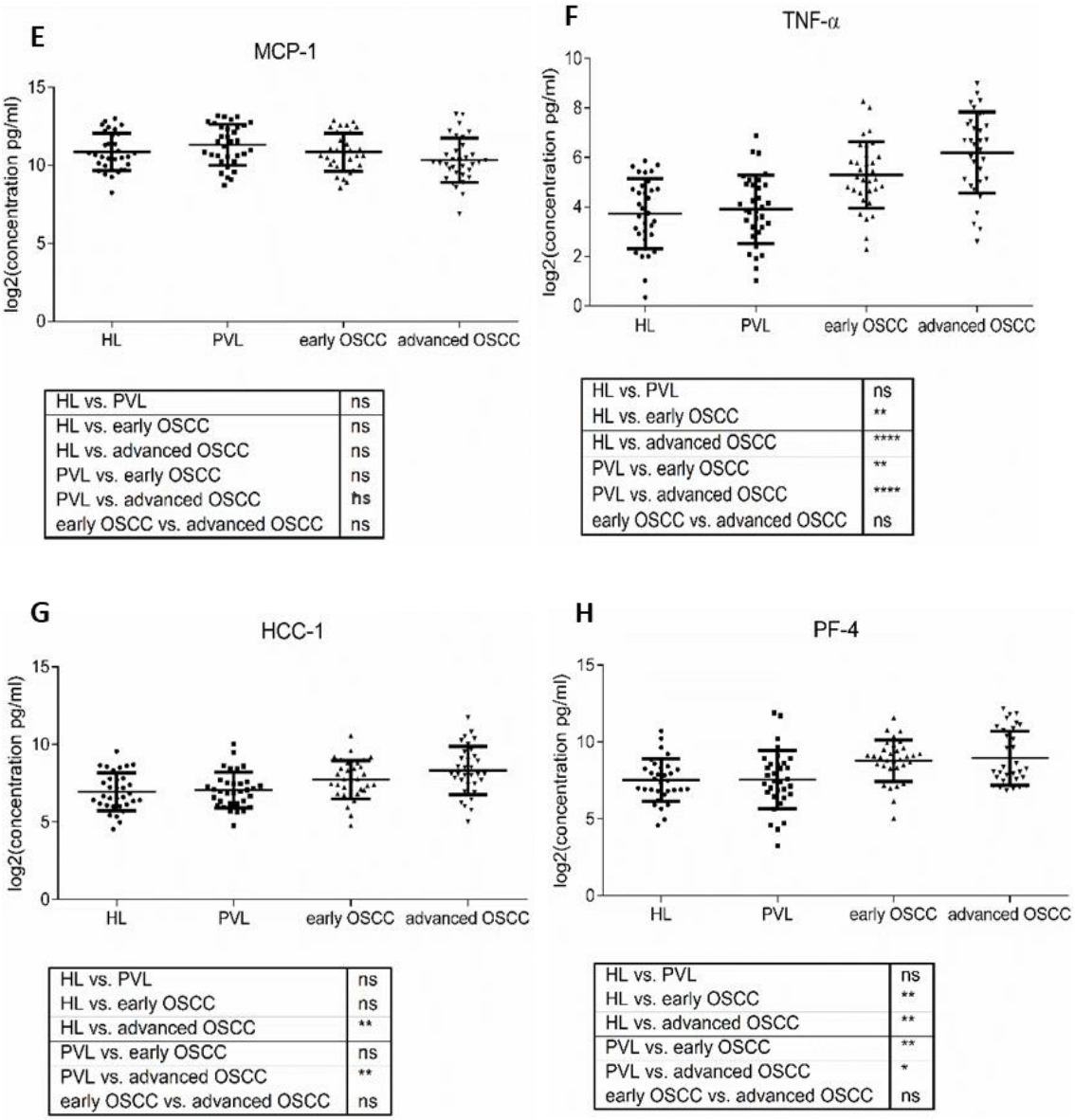


Figure 34. Dot plots showing log2 transformed salivary levels (pg/mL) of (**A**) IL-1α, (**B**) IL-6, (**C**) IL-8, (**D**) IP-10, (**E**) MCP-1, (**F**) TNF-α, (**G**) HCC-1 and (**H**) PF-4 in patients with HL, PVL, early and advanced OSCC stages. Values represent mean ±SEM of n=33 per each group, where n is an average of two technical replicates. Kruskal-Wallis multiple comparisons followed by Dunn's post hoc test were used to estimate statistical significance, considered for a *p*-value less than 0.05 and indicated with **** for p ≤ 0.0001; *** for p ≤ 0.001; ** for p ≤ 0.01 and * for p ≤ 0.05; ns - no significance.

1.5.5 *Correlations among salivary cytokine levels*

Pearson's pairwise correlation test was carried out to find potential correlation among the cytokines within the control and pathology groups. Positive correlation between some analytes was observed in all the groups. In the control cohort (Table 8A, Supplementary tables), IL-1α was positively correlated with IL-8 and TNF-α, with statistically significant correlation (p =0.03 and 0.02, respectively). IL-6 showed association with IL-8, MCP-1, TNF-

α, and HCC-1 (p = 0.002; 0.01; < 0.001 and 0.03, respectively). IL-8 with HCC-1 and TNF-α (p = 0.01; < 0.001, respectively). A relationship was also found between TNF-α and HCC-1 (p = 0.02), as well among HCC-1 and PF-4 (p < 0.001). In the HL group (Table 8B, Supplementary tables), IL-6 exhibited notable relations with IL-8, TNF-α, HCC-1, and PF-4 (p = 0.007; < 0.001; < 0.001 and 0.03, respectively). IL-8 was correlated with IP-10, MCP-1, TNF-α, HCC-1 and PF-4 (p = 0.009; 0.007; < 0.001; 0.01 and 0.03, respectively). IP-10 was associated with PF-4 and HCC-1 (p = 0.006; < 0.001, respectively). Considerable relation was found between MCP-1 and TNF-α (p = 0.003), TNF-α and HCC-1 (p < 0.001), and HCC-1 and PF-4 (p = 0.003). Within the cohort of PVL patients (Table 8C, Supplementary tables), correlation was observed between IL-1α and IL-6, IL-8, MCP-1 and HCC-1 with significance of p = 0.002; < 0.001; 0.03; 0.003, respectively. IL-6 demonstrated considerable relation with IL-8, MCP-1, HCC-1 and PF-4 (p = 0.004; 0.02; 0.005; 0.01, respectively), while IL-8 with IP-10, MCP1-1 and HCC-1 (p =0.004; 0.003; < 0.001, respectively). IP-10 was positively correlated with MCP-1 and HCC-1 (p = 0.02; 0.03, respectively), MCP-1 with HCC-1 and PF-4 (p < 0.001, for both) and TNF-α with HCC-1 and PF-4 (p = 0.008 and 0.02, respectively). Across the early OSCC sample set (Table 8D, Supplementary tables), significantly related were IL-1α with MCP-1 (p= 0.004), and TNF-α (p = 0.002), also IL-6 with IL-8 and TNF-α (p = 0.005 and 0.008, respectively). A connection between elevated levels of IL-8 with MCP-1, TNF-α, HCC-1 and PF-4 was established with TNF-α at p = 0.002; < 0.001; 0.002 and 0.05, respectively), between MCP-1 and TNF-α (p = 0.02) and among TNF-α and HCC-1 (p < 0.001). Within the advanced OSCC group (Table 8E, Supplementary tables), remarkable correlation was seen in IL-6 with IP-10 and TNF-α (p = 0.007 and < 0.001, respectively), in IL-8 with MCP-1, HCC-1 and PF-4 (p < 0.001, for all), in MCP-1 with PF-4 and HCC-1 (p = 0.003 and < 0.001), and lastly among HCC-1 and PF-4 (p = 0.03).

1.5.6 *Multinomial regression analysis*

To assess the potential association of altered cytokine expression with the categories (control, HL, PVL early, and advanced OSCC), multinomial logistic regression analysis was performed. The groups were used as a response variable, considering the control as the reference category, while the mean values of the eight cytokines were used as dependent covariates. A significance table for regression has been generated, including all the analytes (Table 9A). As it can be seen, the Sig column indicates the association of each cytokine,

being statistically significant for one or more pathologies IL-1-α, IL-6, MCP-1, TNF-α, and PF-4. Regression model parameters were recalculated only with the analytes, significantly associated with the response variable (Table 9B). After recalculation of the p-values, the results support the statistical significance of the initially associated cytokines.

Table 9A. Maximum likelihood estimation (MLE) of the influence of each of the eight covariates (cytokines) on a variable categorical response of multiple categories (control, HL, PVL early, and advanced OSCC groups).

Effect	Model selection criteria	Maximum likelihood		
	-2 log-likelihood of the reduced model	Chi-squared	groups	Sig.
Intersection	303.476	5.201	4	.267
IL-1α*	325.974	27.699	4	.000
IL-6*	310.432	12.157	4	.016
IL-8	300.682	2.406	4	.661
IP-10	308.376	10.100	4	.059
MCP-1*	320.364	22.089	4	.000
TNF-α*	324.551	26.275	4	.000
HCC-1	302.794	4.519	4	.340
PF4*	310.172	11.896	4	.018

The chi-squared statistic is the difference in the -2 log-likelihoods between the final model and the reduced model (The -2 log and the chi-square indicate the goodness of the model by including each of the analytes). The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all the parameters of this effect are 0. (significance * is set to $p < 0.05$).

Table 9B. Recalculated MLE with the statistically significant cytokines (IL-1-α, IL-6, MCP-1, TNF-α, and PF-4).

Effect	**Model selection criteria**	**Maximum likelihood**		
	-2 log-likelihood of the reduced model	Chi-squared	groups	Sig.
Intersection	318.749	4.827	4	.306
IL1-α	338.918	24.997	4	.000
IL-6	330.529	16.607	4	.002
MCP-1	344.466	30.544	4	.000
TNF-α	347.971	34.049	4	.000
PF-4	327.255	13.334	4	.010

Regression analysis, shown in Table 10, reveals a potential correlation of IL-1α, IL-6, MCP-1, TNF-α, and PF-4 with a concrete pathological group. Statistically, a significant association was noted between HL and PVL groups and IL-1α, TNF-α, and MCP-1, as well as between PF-4 and PVL groups. Regarding the OSCC groups, an association is observed between both early and advanced stages with IL-1α and TNF-α, while IL-6 is likely related to advanced OSCC patients.

Table 10. Summary of multinomial logistic regression analysis showing the significance of IL-1α, IL-6, MCP-1, TNF-α, HCC-1, and PF-4 in association with the pathology groups (HL, PVL, early and advanced OSCC).

class		B	St E	Sig.	Exp(B)	Confidence interval at 95% for Exp(B)	
						Inferior limit	Superior limit
Advanced OSCC	**Intersection**	**4.611**	**4.623**	**.319**			
	IL-1α	-1.504	.397	.000	.222	.102	.484
	IL-6	1.208	.455	.008	3.347	1.373	8.159
	MCP-1	-.670	.456	.142	.512	.209	1.252
	TNF-α	2.522	.572	.000	12.459	4.064	38.196
	PF-4	.182	.386	.638	1.200	.562	2.559
Early OSCC	**Intersection**	**1.720**	**4.211**	**.683**			
	IL-1α	-1.346	.348	.000	.260	.132	.515
	IL6	.845	.434	.051	2.327	.995	5.444
	MCP1	.117	.384	.761	1.124	.529	2.387
	TNF-α	2.004	.529	.000	7.420	2.630	20.930
	PF-4	-.019	.353	.957	.981	.491	1.961
HL	**Intersection**	**.413**	**3.876**	**.915**			
	IL-1α	-.708	.305	.020	.493	.271	.896
	IL-6	.486	.415	.242	1.625	.721	3.666
	MCP-1	.654	.325	.044	1.923	1.017	3.635
	TNF-α	.958	.472	.042	2.606	1.033	6.575
	PF-4	-.568	.303	.060	.567	.313	1.025
PVL	**Intersection**	**-3.022**	**3.998**	**.450**			
	IL-1α	-.694	.316	.028	.499	.269	.928
	IL-6	.159	.428	.710	1.172	.507	2.712
	MCP-1	1.048	.337	.002	2.851	1.473	5.522
	TNF-α	1.093	.479	.023	2.983	1.167	7.626
	PF-4	-.617	.308	.045	.540	.295	.987

Multinomial regression analysis informs about the implication (measured by the coefficients B) of each of the analysed variables (cytokines) in the category of interest (studied group). The reference category is the control group; B = estimation of the regression coefficients, St E = Standard error, Sig = statistical significance at 95% confidence;p <0.05. In orange are marked significant cytokines according to the multiple regression analysis (MRA). In red are those, significant in both the MRA test and in the comparisons between the pathology groups against the control, shown in chapter 1.5.

1.5.7 *Association of cytokine levels with patients' clinical variables*

For these analyses, a Bayesian model has been applied using directed acyclic graphing (DAG) and conditional probability methodology. DAGs have been constructed to establish existing associations between altered expression of the studied cytokines in the pathology groups (HL, PVL, early and advanced OSCC) and patient's clinical variables.

1.5.8 *Correlation between altered cytokine expression and clinical variables of patients with HL and PVL*

To estimate the potential correlation between altered cytokine levels and clinical parameters of patients with HL and PVL, the following variables have been taken into consideration:

- sex (two conditions: males and females)
- smoking habits (three conditions: ex, current, and non-smokers)
- oral location (6 conditions: lip, buccal mucosa, gingiva, tongue, the floor of the mouth, palate)
- lesion clinical type (three conditions: homogeneous white plaques, verrucous lesions, and mixed type)
- histologic features (two conditions: the presence of epithelial dysplasia (ED) and no ED
- salivary levels of IL-1α, IL-6, IL-8, IP-10, MCP-1, TNF-α, HCC-1, and PF-4.

The positive correlation found between cytokine concentration and clinical variables of HL and PVL patients is visualized in Figure 35. A relationship was observed between histologic features of premalignant lesions and HCC-1, which could be used to indicate the presence

of epithelial dysplasia (ED). IL-6 was correlated with HCC-1, thus it might be considered as an associated marker. However, probability values of HCC-1 and IL-6 to determine dysplasia remain close (Table 9A and B, Supplementary tables). TNF-α, PF-4, IL-6, IP-10, and IL-8 showed association with lesions clinical type, though an inter-cytokine correlation was also observed, suggesting a mutual influence on the expression levels. Information regarding the concentration ranges that maximizes the probability to associate the analytes with mixed, verrucous, and homogenous lesions is shown in Table 9 C, D, and E, respectively (Supplementary Tables). No correlation was found between sex, smoking habits, lesion oral location, and cytokine concentrations.

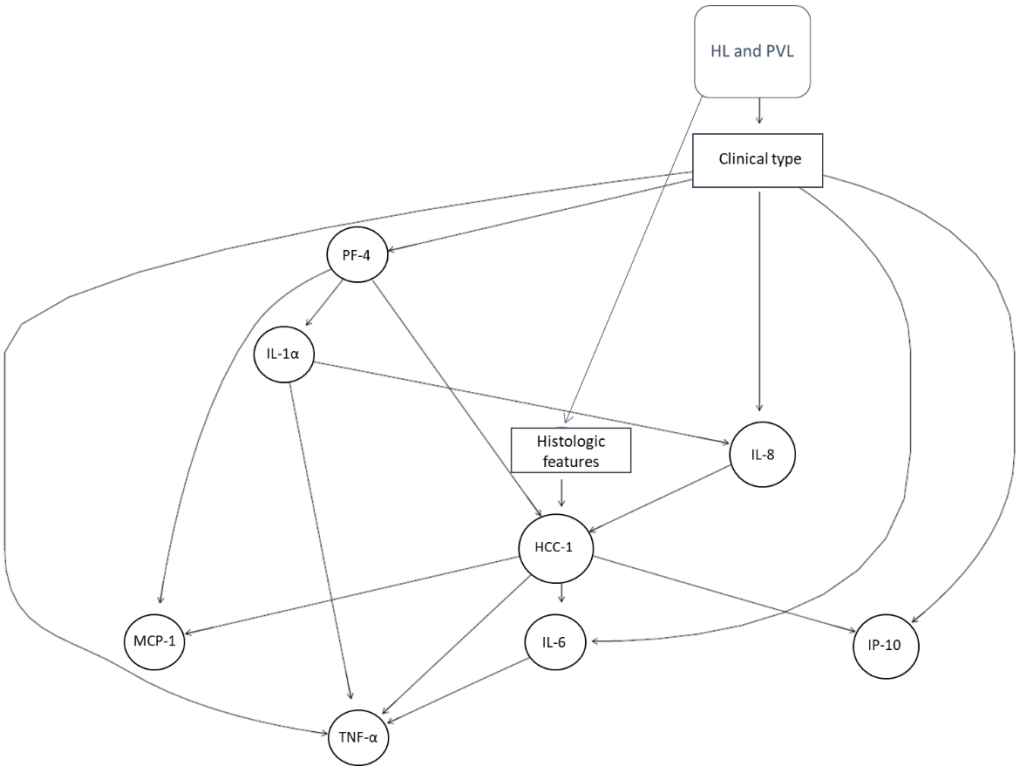


Figure 35. DAG graph representing the correlation of altered cytokine expression with lesion clinical type and histologic features of patients diagnosed with HL and PVL. Clinical parameters for which correlation was not estimated are omitted from the graph.

1.5.9 *Correlation between altered cytokine expression and clinical variables of OSCC patients*

For correlation analysis of altered cytokine levels and clinical parameters of patients diagnosed at early and advanced OSCC stages, the following variables have been taken into account:

- sex (two conditions: males and females)

- smoking habits (three conditions: ex, current, and non-smokers)
- oral location (6 conditions: lip, buccal mucosa, gingiva, tongue, the floor of the mouth, palate)
- lesion clinical type (three conditions: ulcerative, exophytic, and mixed)
- lesion size (3 conditions: 0-2 cm;2-4 cm; >4 cm)
- cervical adenopathy (2 conditions: yes and no)
- histologic features (two conditions: well and not well-differentiated tumors)
- expression levels of IL-1α, IL-6, IL-8, IP-10, MCP-1, TNF-α, HCC-1, and PF-4.

Visual representation of correlations found between modulated salivary cytokine levels and OSCC patient's clinical variables is shown in Figure 36. A direct relation was found between OSCC and TNF-α. Further, a direct association of OSCC histologic features was observed with HCC-1, MCP-1, and PF-4 and indirect with TNF-α which could be used to distinguish well from not well-differentiated tumors with relatively high probability (0.6 - 0.8) considering the value ranges of the analytes (Table 10A and B, Supplementary tables). The presence of cervical adenopathy was positively correlated to IL-6 levels, which itself appeared associated with IP- 10. According to the concentration ranges of IL-6 and IP-10, the probability to signify the existence of cervical adenopathy in OSCC was shown to reach up to 0.8 (Table 10 C and D, Supplementary tables. No relation was found among sex, smoking habits, tumour oral location, clinical type, and size, and altered cytokine expression.

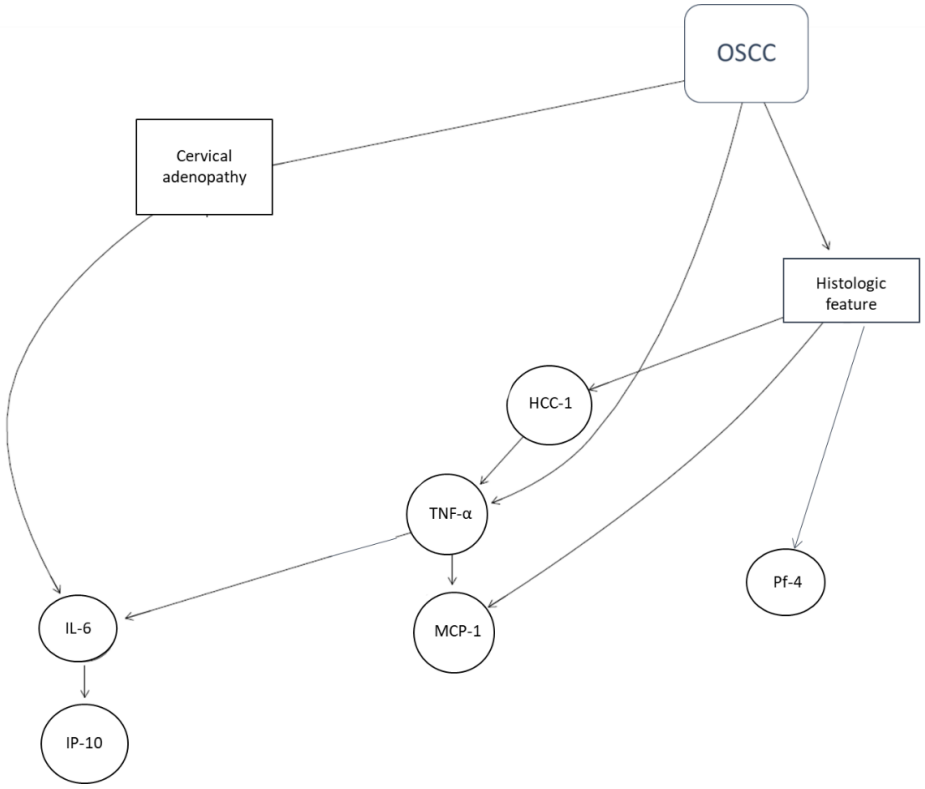


Figure 36. DAG graph representing the positive association between altered cytokine expression, histologic features, and the presence of cervical adenopathy in patients at early and advanced OSCC. Clinical parameters for which correlation was not estimated are omitted from the graph.

1.6 *Diagnostic potential of salivary cytokines*

To study the diagnostic potential of the eight cytokines, DAG methodology was utilized, integrating the following variables:

- patient cohorts (with five states: Control, HL, PVL, early OSCC, and advanced OSCC)
- expression levels of IL-1α, IL-6, IL-8, IP-10, MCP-1, TNF-α, HCC-1 and PF-4

A Bayesian network using DAG methodology and conditional probability was implicated to establish an association between the pathology groups and the studied cytokines (Fig.37). The analysis showed a direct association of MCP-1, HCC-1, TNF-α, IL-6, IL-8, PF-4, and IP-10 with the groups. Also, it can be seen that proteins are influenced by each other, such as TNF-α which appears affected by IL-6 but itself impacts HCC-1 and IL-8 levels. Thus, IL-6 is not influenced by the other analytes is suggested as a potential independent diagnostic marker with the probability of association with OSCC reaches up to 70% (Table 11E, Supplementary Tables). TNF-α, IL-8, PF-4, HCC-1, and IP-10 could be considered as associated biomarkers due to the established correlation with IL-6. Nevertheless, protein

concentration ranges remain very close, and the probability to differentiate between control, HL, PVL, and early OSCC less than 0.5 (Table 11 A, B, C, and D; Supplementary tables).

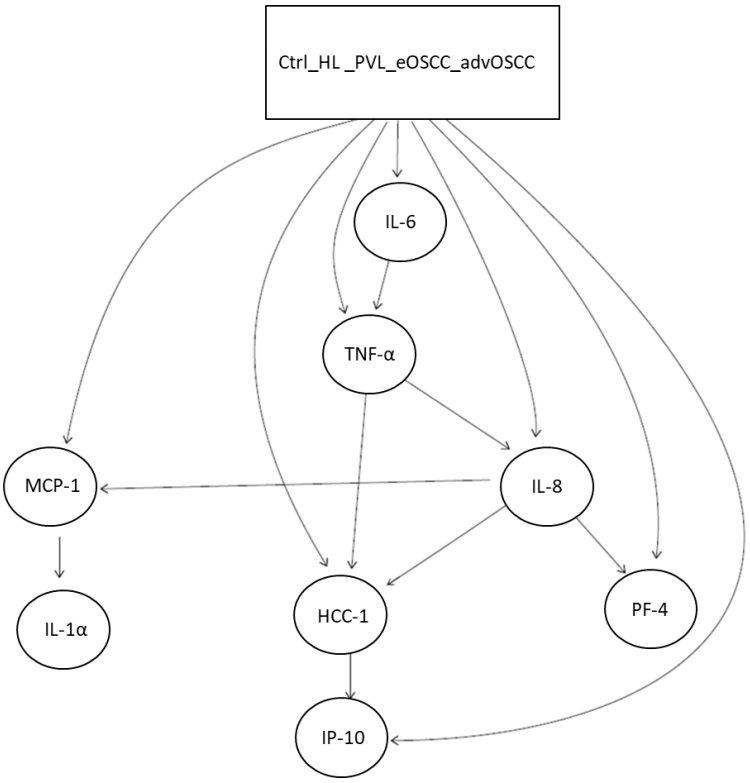


Figure 37. DAG graph showing the correlation between altered cytokine levels and control, HL, PVL early, and advanced OSCC groups.

2. *Salivary N-glycome profiling*

This investigation was performed to describe and compare *N*-glycome profiles of patients with premalignant lesions (PVL), early and advanced OSCC stages, and age-matched controls.

2.1 *Participants' features*

For *N*-glycan profiling, saliva samples were collected from twenty representatives among the recruited project participants and equally distributed into four groups according to the established case-control inclusion criteria (Table 12). Group 1 consisted of five PVL patients (mean age 67 years), Group 2 (mean age 74) and 3 (mean age 60) comprised 5 cases each,

including patients at early (I and II) and advanced (III and IV) OSCC stages, respectively. The Control - group 4 (mean age 55) included five healthy subjects, without visible oral lesions.

Table 12. Participant information summary

	Controls (n=5)	PVL (n=5)	Early OSCC (n=5)	Advanced OSCC (n=5)
Age (median ± SD)	55 ± 9.8	67 ± 10.6	74 ± 4.5	60 ± 8.8
Sex				
Male	1	2	1	3
Female	4	3	4	2

2.2 *Identification and initial description of salivary N-glycans*

To clarify the glycan profiles of individuals with and without oral lesions, total *N*-glycans released from pooled salivary proteins were identified. Primary HILIC-MS sample analysis of controls, patients with PVL, early and advanced OSCC stages disclosed complex profiles with multiple peaks annotation, even after HILIC separation. MS data acquisition was crucial in performing an exhaustive characterization, revealing 90 compositions, firstly identified in the total salivary pool of the controls and subsequently in the other three groups. Next-generation LC-MS data assessment software (Progenesis QI) enabled statistical analysis of the identified compounds. The selection criteria were based on 1) ANOVA p-value of the mean intensity from nine technical replicates ≤0.05 and 2) *N*-glycan detection in the retention time frame period between 3.5-28 min. Summary including proposed *N*-glycan monosaccharide compositions, their experimental m/z values observed on the MS, theoretical masses, mass accuracy (ppm), and ion charge is shown in Table 13 (Supplementary Tables). All compounds share a common core sugar sequence Manα1–6(Manα1–3) Manβ1–4GlcNAcβ1–4GlcNAcβ1 (206) and due to their labeling with the fluorophore 2-AB, a terminal suffix -core 2AB was added for shortness. Full MS analysis of released *N*-glycans does not allow to perform an accurate determination of monosaccharides linkage positions; as a consequence, several chromatographic peaks with identical MS signals may correspond to isobaric *N*-glycans, described with identical monosaccharide composition.

2.3 *N-glycan profiles of Controls, PVL, early and advanced OSCC stages*

PCA was conducted to estimate the capacity of salivary *N*-glycan profiles to discern between groups. Relative abundance profiles expressed statistical plots of glycan amount, grouped according to Progenesis QI criteria, defined above. The studied groups were distinguished by their first and second principal component scores (PC1 representing the direction of maximum variation and PC2- the highest variation through the data, respectively). The principal component biplot demonstrated segregation of control, PVL, early and advanced OSCC cases (Fig.38). PCA replicate clustering suggests group-specific *N*-glycosylation patterns which differ between control and pathology cases, possibly related to disease-dependent *N*-glycomic aberrations.

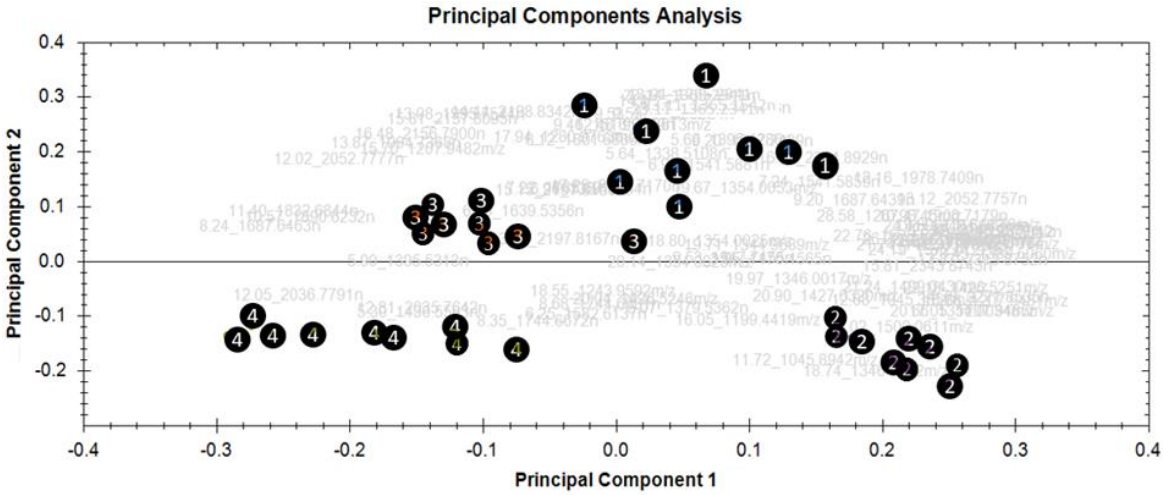


Figure 38. The principal component plot of whole saliva *N*-glycans distribution using integrated HILIC-MS chromatograms of five pooled samples per group: Control - ❶, PVL - ❷, early OSCC - ❸ and advanced OSCC stages- ❹. Three independent sample fractions were evaluated within three technical replicates per group.

Representative UPLC-HILIC-FLC chromatograms of saliva-derived *N*- glycans from control, patients with PVL, and early and advanced OSCC groups displayed complex *N*-glycomic profiles with multiple peaks (Fig.39). Quantitation of annotated peaks is related to the intensity of the MS signal of each glycan eluted at a certain retention time and was utilized to estimate relative abundance differences. To evaluate potential differences of saliva obtained *N*-glycans between healthy subjects, patients with PVL, and different OSCC stages, relative abundance was compared between the groups.

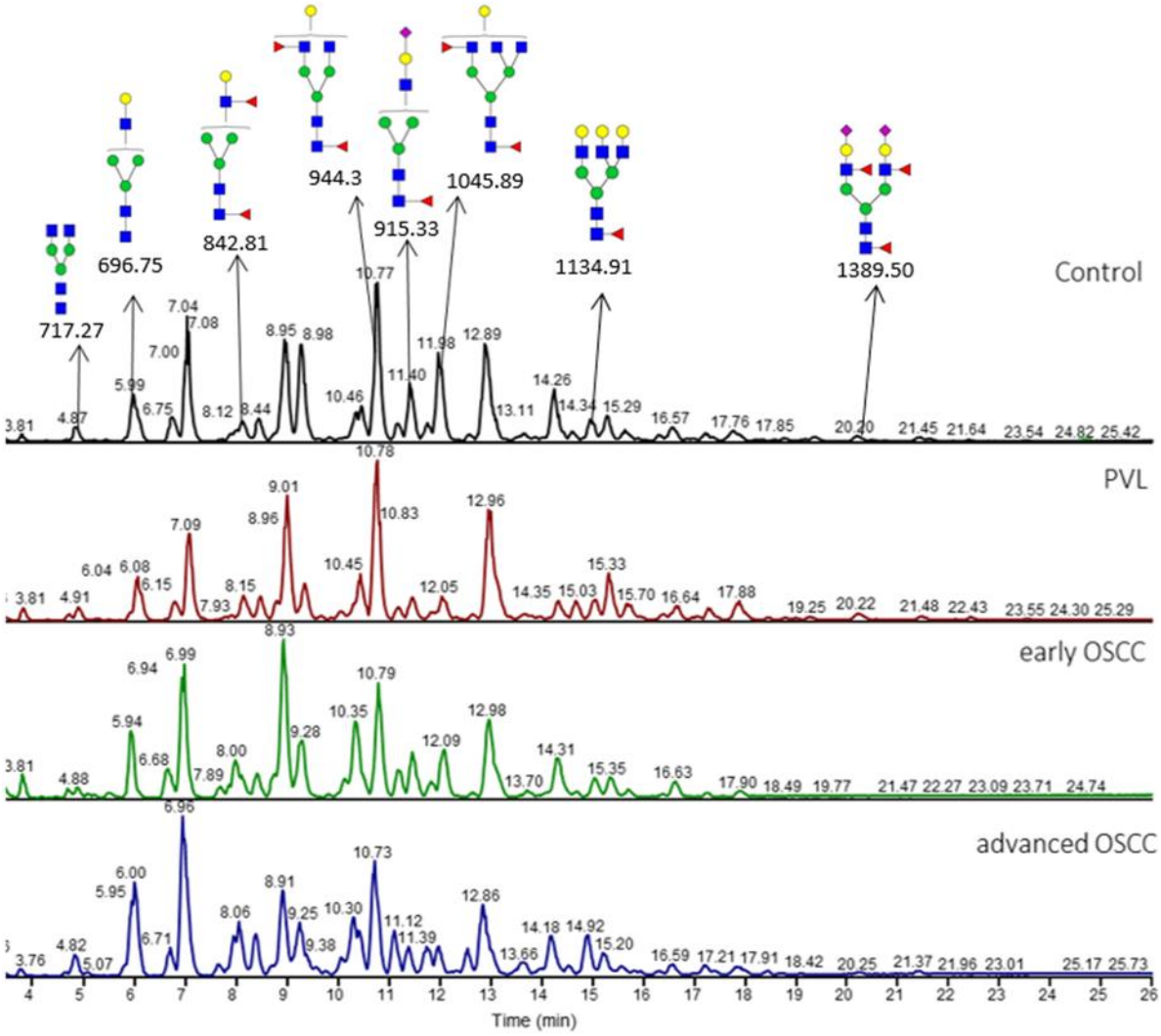


Figure 39. A representative UPLC-HILIC- fluorescent chromatograms of *N*-glycans derived from pooled salivary proteins of control, patients with PVL, early and advanced OSCC stages. Pictograms are representing the eight *N*-glycans exhibiting significantly different relative abundance in OSCC saliva, initially identified in the control pool of samples. Symbols: ■, N-acetylglucosamine; ●, mannose; ○, galactose; ▲, fucose; ◆, N-acetylneuraminic acid (sialic acid).

After further statistical analysis of the annotated compounds, eight *N*-glycans exhibited significantly different relative abundance in OSSC compared to PVL and controls (Fig.39). Three out of eight demonstrated decreased relative abundance in OSCC patient saliva compared with PVL and Control. These glycans included core and antennary fucosylated tri-antennary glycan (*m/z*=1045.89, Fig 40A), core fucosylated tri-antennary glycan (*m/z*=1134.91, Fig 40B) and core and antennary fucosylated di-sialylated bi-antennary glycan (*m/z*=1389.50, Fig 40C).

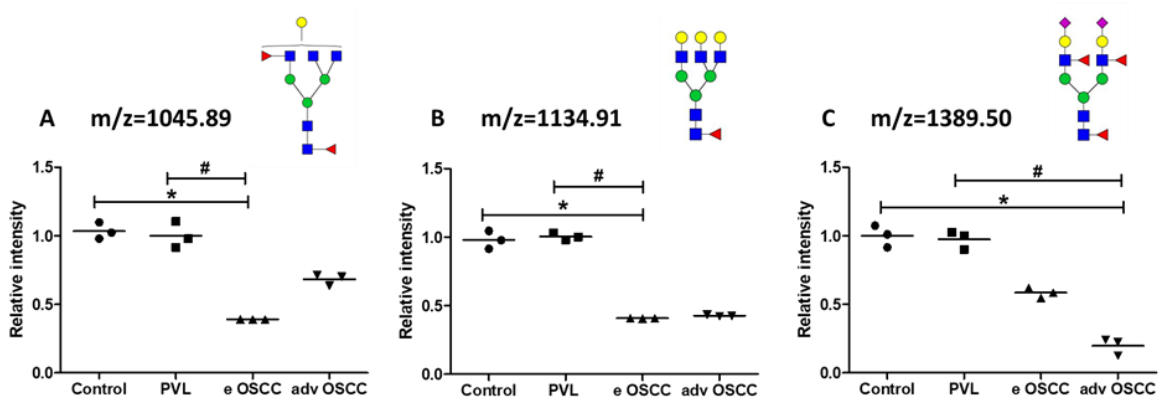
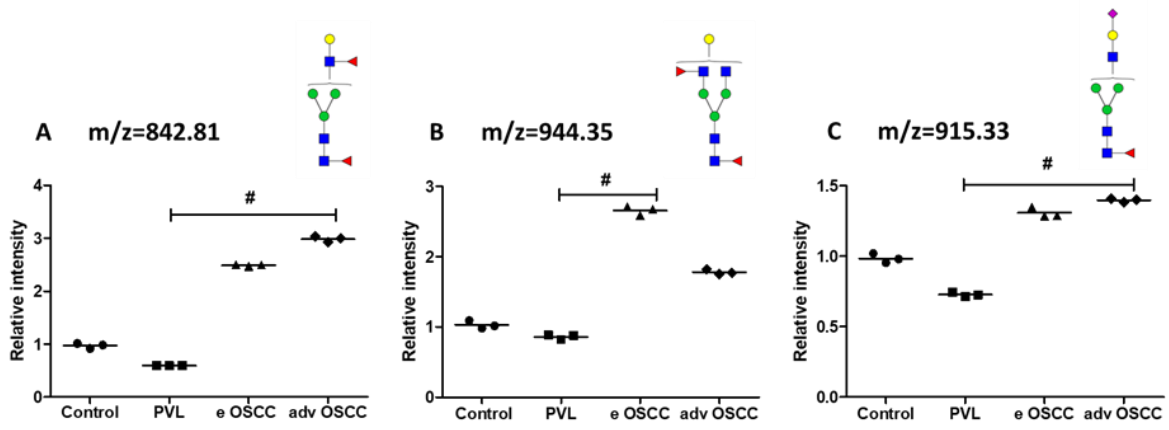


Figure 40. Scatter diagrams of the three *N*-glycans showed decreased relative abundance in early (e) and advanced (adv) OSCC compared to PVL and controls. The dot plot of the relative abundance of (**A**) core and antennary fucosylated tri-antennary glycan (*m/z*=1045.89), (**B**) core fucosylated tri-antennary glycan (*m/z*=1134.91), and (**C**) core and antennary fucosylated di-sialylated bi-antennary glycan (*m/z*=1389.50) in saliva. Abundance expression is presented as fold-change relative to Control (Control = 1). Difference between early and advanced OSCC versus PVL and Control groups, analysed by nonparametric Kruskal Wallis test where * $P < 0.05$ OSCCversus Control and $^{\#}P < 0.05$ versus PVL. Data are shown as the mean of three technical replicates per group, each one consisted of 5 pooled samples (n=5).

Among the eight *N*-glycans found to have considerably altered expression between control and pathology cases, five exhibited increased relative abundance in early as well as in advanced OSCC saliva. The dot plots representing their relative abundance were shown in Figure 41. The five *N*-glycans included core and antenna fucosylated/bi-fucosylated *N*-linked glycan (*m/z*=842.81, Fig. 41A), core and antenna fucosylated/bi-fucosylated bi-antennary glycan (*m/z*=944.35, Fig.41B), mono sialylated core fucosylated glycan (*m/z*=915.33, Fig.41C), bi-antennary *N*-linked glycan (*m/z*=717.27, Fig.41D) and *N*-linked glycan (*m/z*=696.75, Fig.41E).



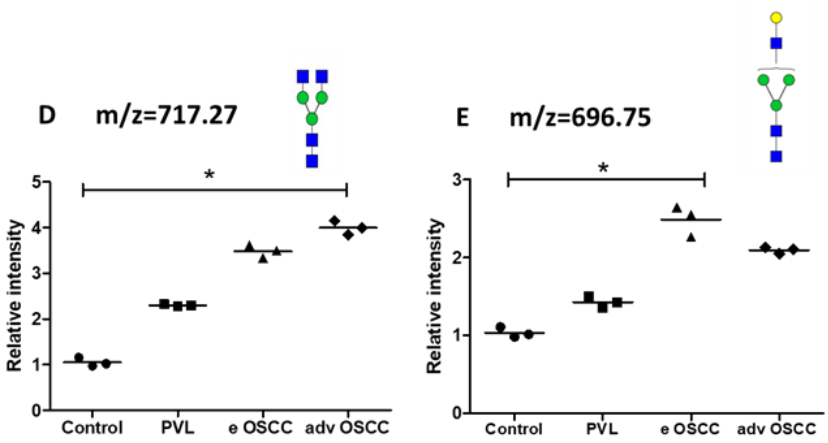


Figure 41. Scatter diagrams of *N*-glycans showed increased relative abundance in the groups of patients with PVL, early (e), and advanced (adv) OSCC stages compared with Controls. The dot plot of the relative abundance of (**A**) core and antenna fucosylated glycan (observed at *m/z*=842.81), (**B**) core and antenna fucosylated bi-antennary glycan (observed at *m/z*=944.35), (**C**) mono sialylated core fucosylated glycan (observed at *m/z*=915.3), (**D**) bi-antennary glycan (observed at *m/z*=717.27) and (**E**) *N*-linked glycan (observed at *m/z*=696.75). Abundance expression is presented as fold-change relative to Control (Control = 1). Data are shown as the mean of three technical replicates per group, each one consisted of 5 pooled samples (n=5). * $P < 0.05$ vs Control and $^{\#}P < 0.05$ vs. PVL.

3. *Salivary proteome profiling*

LC-Mass spectrometry (MS/MS) analysis was carried out for qualitative and quantitative salivary protein profiling of patients with potentially malignant disorder (PVL), at early and advanced OSCC stages, and healthy subjects.

3.1 *Participants demographic and clinical traits*

For the construction of the salivary proteome spectral library, saliva samples from ten patients at different OSCC stages were collected and subjected to LC-MS/MS analysis. A summary of their demographical and clinical traits is shown in Table 14A. The mean age of the OSCC patients was 63 years (range 41-95), including patients at early and advanced clinical stages. For salivary proteome profiling, forty whole saliva samples were selected and assigned into four groups (Table 14B). The median age of the PVL group was 65 years (range: 47-87). The median age of the early OSCC was 67 years (range: 53-85) and of the advanced OSCC group was 74.5 years (range: 58-95) comprised of ten cases each, including patients at I and II, and III and IV clinical stages, respectively. The Control - group also

consisted of ten volunteers, without visible oral lesions the mean age of whose was 58.5 years (range: 48-82).

Table 14. Characteristics of study participants in (**A**) OSCC salivary protein spectral library and (**B**) SWATH analysis and comparison of proteome profiles between patients with PVL, early, advanced OSCC stages, and healthy counterparts (controls), where n=10 per each group.

A	**OSCC**
Age (median±SD)	63 ± 16.5
Sex	
Male	4
Female	6
OSCC stages	
(I and II)	3
(III and IV)	7

B	**Control**	**PVL**	**early OSCC**	**advanced OSCC**
Age (meadian±SD)	58.5 ± 10.8	65 ± 10.1	67 ± 11.55	74.5 ± 11.54
Sex				
Male	5	5	4	6
Female	5	5	6	4

3.2 *Shotgun proteomics analysis of saliva samples*

The direct analysis of complex protein mixtures to generate a global profile of the protein complement within the mixture is referred to as shotgun proteomics. The shotgun approach was based on nano-LC separation followed by data-dependent acquisition (DDA) using nano ESI-Q-TOF MS. The obtained MS/MS spectra were automatically processed and submitted for search in the human protein sequence database for protein identification. Although there may be oral bacterial proteins present in the human whole saliva, due to the focus of the study to recognize only human proteins, identification searches were limited to the human species database. A peak list was generated directly from the wiff files (MS data format) (Fig.42). The Paragon algorithm (220) of ProteinPilot version 5.0 was used to search the Swissprot database (version 03-2018) with the following parameters: trypsin specificity, cys-alkylation, taxonomy restricted to human, and the search effort set to through. The protein grouping was done by the Pro group algorithm: a protein group in a Pro Group Report is a set of proteins that share some physical evidence. Unlike sequence alignment analyses where full-length theoretical sequences are compared, the formation of protein groups in the Pro Group was guided entirely by observed peptides only. Since the observed peptides are determined from experimentally acquired spectra, the grouping can be considered as guided by the usage of spectra.

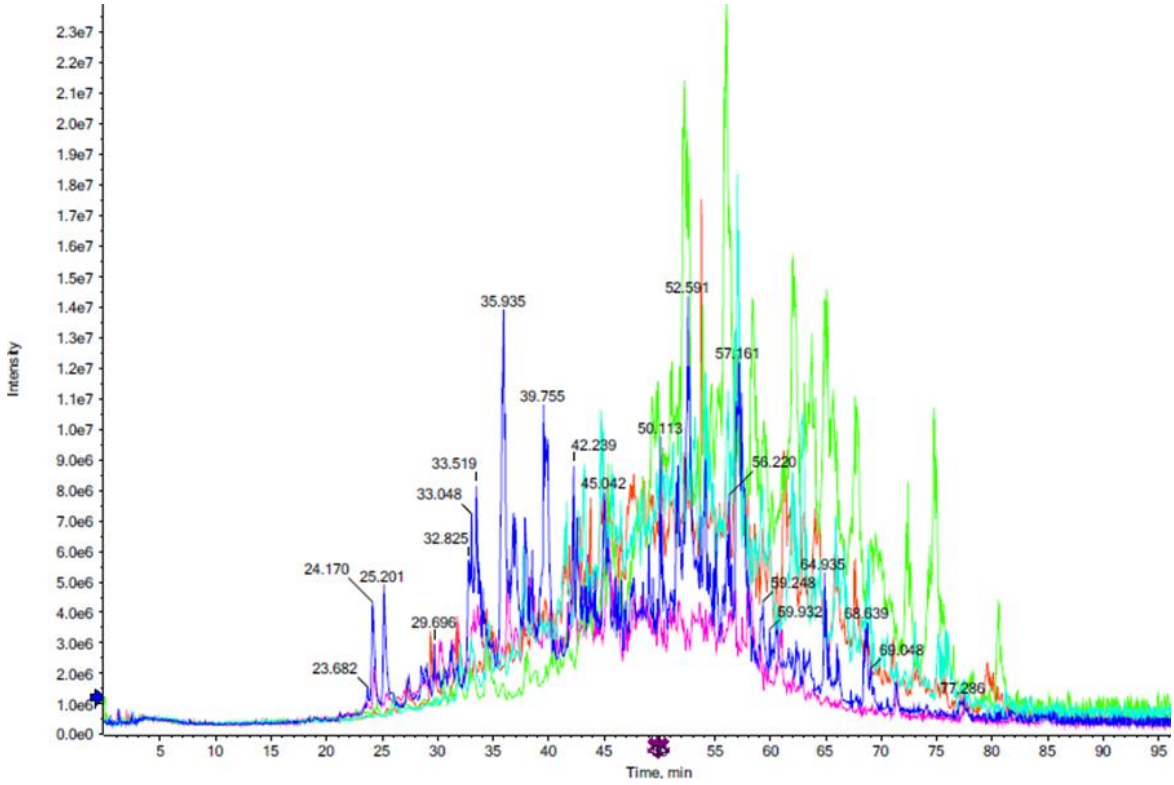


Figure 42. Total Ion chromatogram (TIC) of ionized proteins in pooled OSCC saliva samples.

Typically, in proteomics, a measure for characterizing the performance of a method is the computation of false discovery rate (FDR). It is the metric for the confidence assessment of a large-scale proteomics dataset. ProtScore is another measure of confidence for a detected protein, calculated from the peptide confidence for peptides from spectra that are not already completely "used" by higher-scoring winning proteins. For building the OSCC spectral library were considered proteins that showed Protscore units >1.3 corresponding to ≥ 95% of identification confidence and with 1% Global FDR Fit. After removing the redundant assignments, in total 1053 proteins have been identified (Table 15) in the total salivary pool of OSCC whole saliva.

Table 15. Number of proteins identified at critical false discovery rate types in OSCC saliva

Critical FDR	Local FDR	*Global FDR*	Global FDR from Fit
1.0%	*817*	*891*	**1053**
5.0%	**931**	*1284*	*1259*
10.0%	**1065**	*1408*	*1379*

FDR- false discovery rate**; Local FDR** –measures the FDR of an individual ID; **Global FDR** measures the FDR of a collection of IDs

Out of the 1053 proteins identified, 691 proteins (76%) were matched with two or more peptides. However, the other 234 (24%) were identified based on single hits. If a single peptide satisfied the criteria of Swissprot search, then it was assumed that its original protein existed in the fraction.

3.3 *Salivary proteome profiles of controls, patients with PVL, and OSCC stages*

To discover alterations of proteome profiles, SWATH analysis was used for quantification of proteins in 40 saliva samples: 10 PVL, 10 early and 10 advanced OSCC patients and 10 control subjects without oral lesions. These patients and healthy individuals were matched in terms of sex, ethnicity, and age to minimize potential variation from these factors during the discovery phase. Protein samples were initially loaded into SDS/PAGE and subsequent in-solution tryptic digestion with LC-MS/MS of the resulting peptides allowed the identification of proteins in each collected fraction. The retention times were aligned among the different samples using main protein peptides. The chromatogram of the fragment ions stored in the spectral library was extracted and their peak areas were determined. SWATH generated information regarding the relative abundance of 691 proteins, overlapping between disease and control. Annotation including UniProt protein ID, name and logarithmically transformed mean of 10 individual samples (log2Med) per each group (control, PVL, early and advanced OSCC) and main subcellular location are shown in Table 16 (Supplementary Tables). Figure 43 represents quantitative data distribution, showing similar dispersion across the individual samples in each group. Out of the 691 proteins (FDR <1%), 340 (49.2%) were quantified with three or more peptides, while 351 (50.8 %) were quantified based on one or two peptide matches.

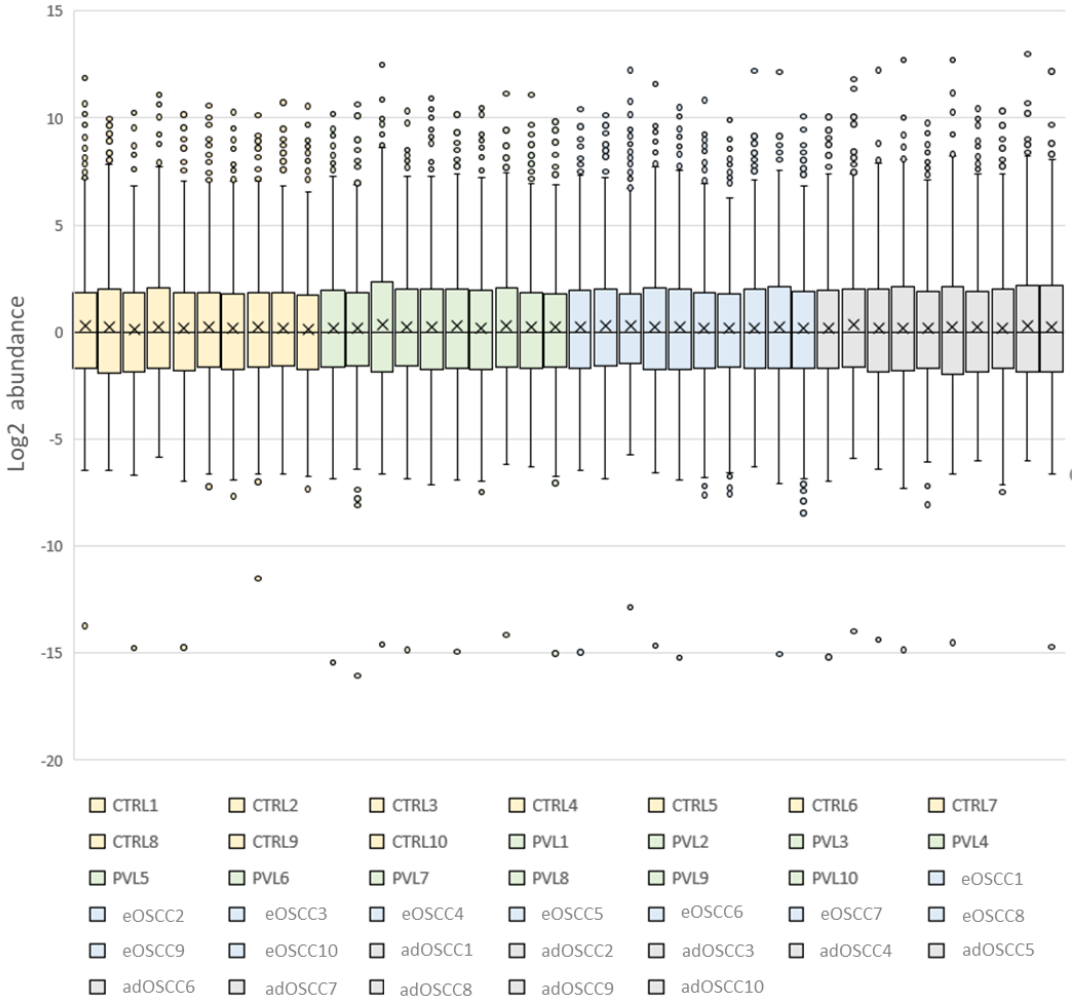


Figure 43. Box plot showing the distribution of normalized quantitative protein data across individual saliva samples of PVL, early and advanced OSCC (eOSCC and adOSCC, respectively), and control (ctrl) cohorts; n=10 per each group.

To evaluate the capability of salivary proteome profiles for disease grouping, statistical tests of reduction of the dimensionality such as principal component analysis (PCA) and discriminant analysis DA (both with Pareto scaling) were carried out. Integration data for 40 samples were aligned and subjected to multivariate PCA analysis, summarizing the overall variability among individuals. The groups investigated, control, PVL, early and advanced OSCC, are distinguished by their first and second principal component scores (PC1 and PC2, respectively). The PC plot showed partial segregation of control and pathologies profiles (Fig.44).

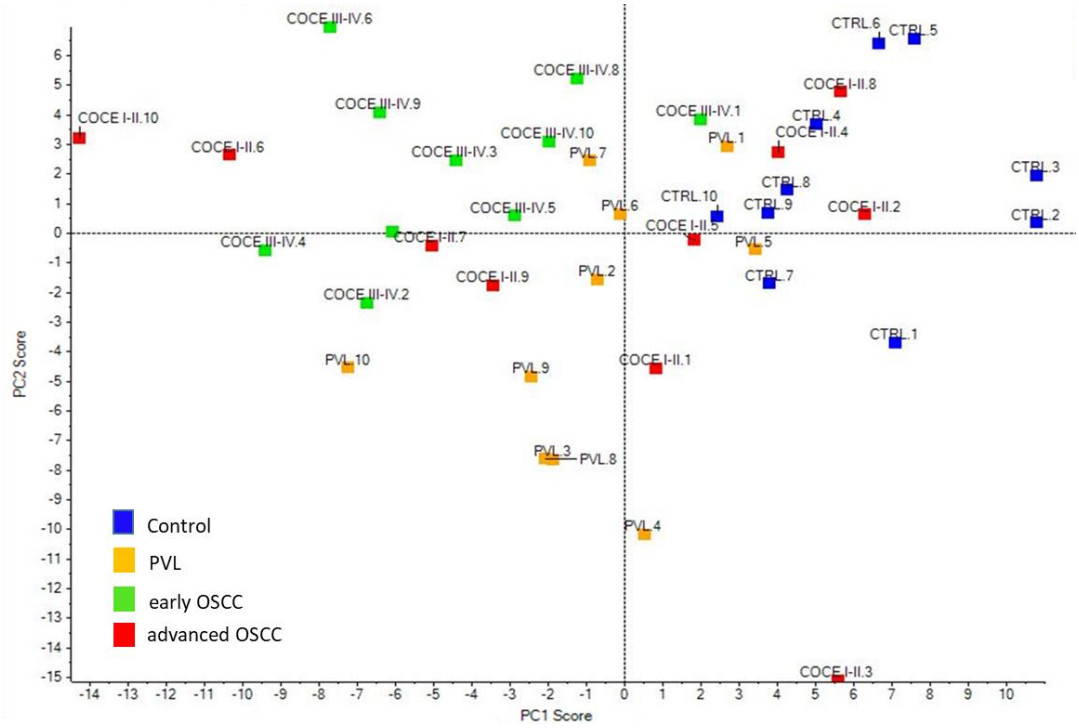


Figure 44. PCA plot of whole saliva proteomes profiles of 10 individuals per Control, PVL, early and advanced OSCC group, based on first and second PC scores.

DA relies on data transformation using PCA as a prior step, which ensures that variables (in this case, the studied groups) submitted to DA are uncorrelated and that their number is less than that of analysed individuals. This multivariate method was used to identify clusters of related samples, overlooking within the defined groups. Observed clustering of control, PVL, early and advanced OSCC proteome profiles indicates differentiation and variability between the investigated groups (Fig.45).

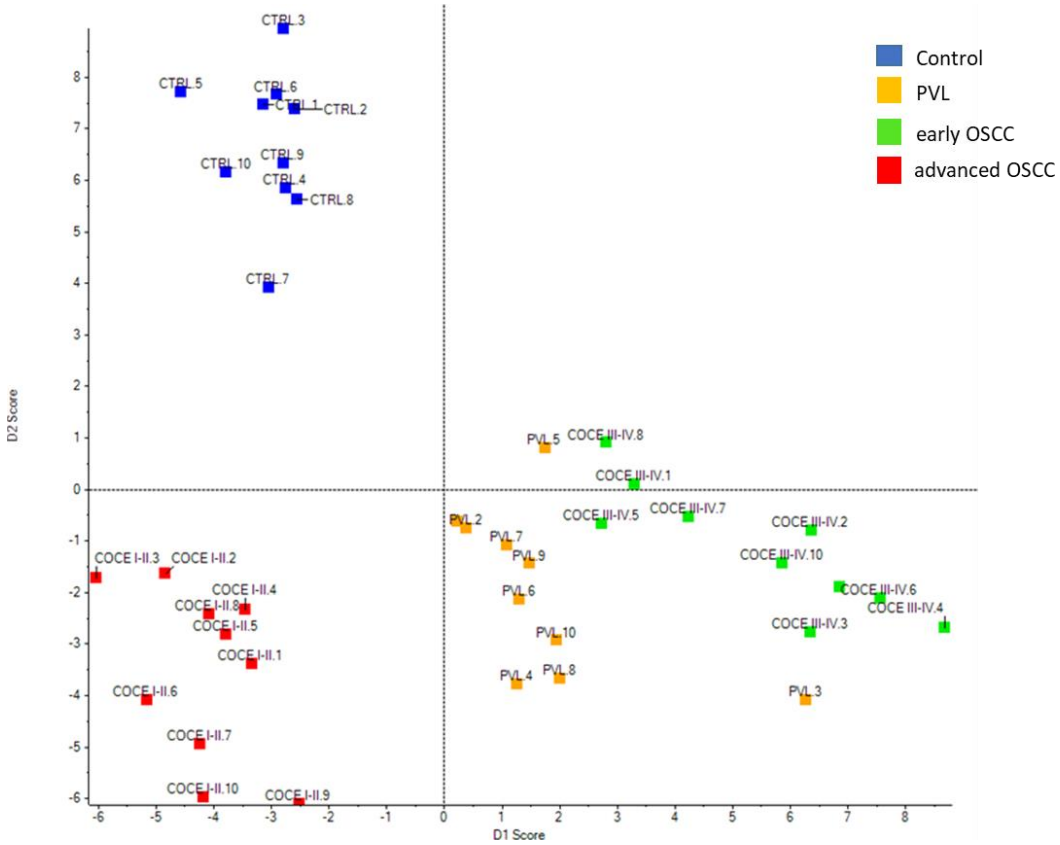


Figure 45. Discriminant analysis of whole saliva proteome profiles belonging to 10 PVL, 10 early, 10 advanced OSCC patients, and 10 healthy controls based on the first and second discriminant scores.

3.4 *Differentially expressed proteins*

To determine what proteins have the potential to differentiate PVL, early, advanced OSCC, and control groups, quantitative data were logarithmically transformed and z-scores were calculated. The heat map in Fig.46 displays columns clustering individual samples and rows quantified proteins, both represented in an order according to a hierarchical classification with standardized values, to avoid scaling issues. Differences are observed at protein level, but they cannot be associated with the investigated groups. Differential proteins among the four groups are listed in Table 17 (Supplementary tables)

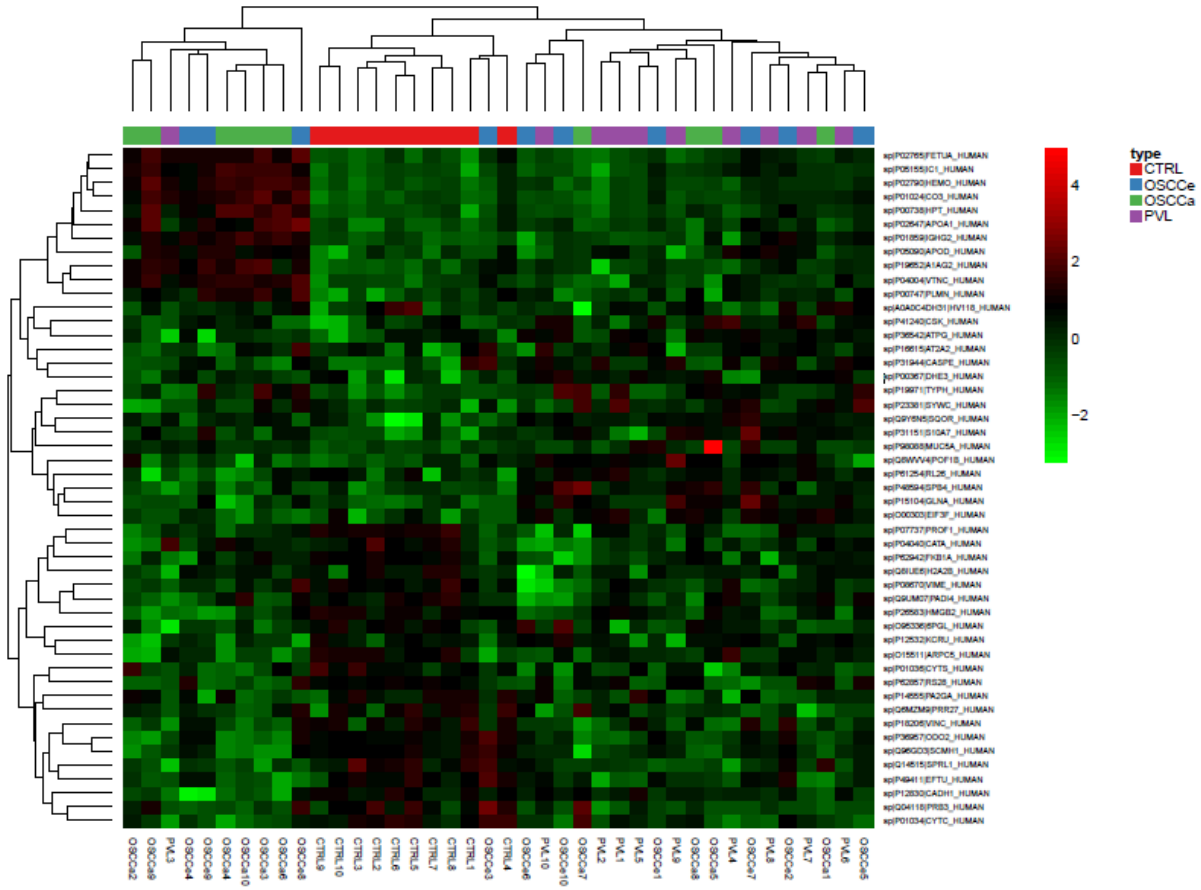


Figure 46. Cluster analysis and heat map of quantified proteins in 10 saliva samples from patients with PVL, different OSCC clinical stages, and control subjects. Proteins are clustered in rows and samples in columns. The colour key at the top of the heat map shows group clumps as PVL- purple; early (e) OSCC – blue; advanced (a) OSCC- green and control (ctrl) - red. The scale at the right top shows log2 normalized protein abundance and ranges from low (green) to high (red) levels, according to the Z-score value. The heat map scale of Z scores ranges from -4.5 (green) to 4.5 (red) with a midpoint of 0.0 (black).

Partial least squares discriminant analysis (PLS-DA) enables the selection of the most predictive and discriminative features in the data that allow categorization of samples using a linear classification model. A PLS-DA classification plot was generated, using quantitative data of the annotated proteins per each of the studied groups. In Fig.47 it can be observed segregation of control from pathology cases, however PVL, early and advanced OSCC samples did not show clear distribution according to the group of belonging.

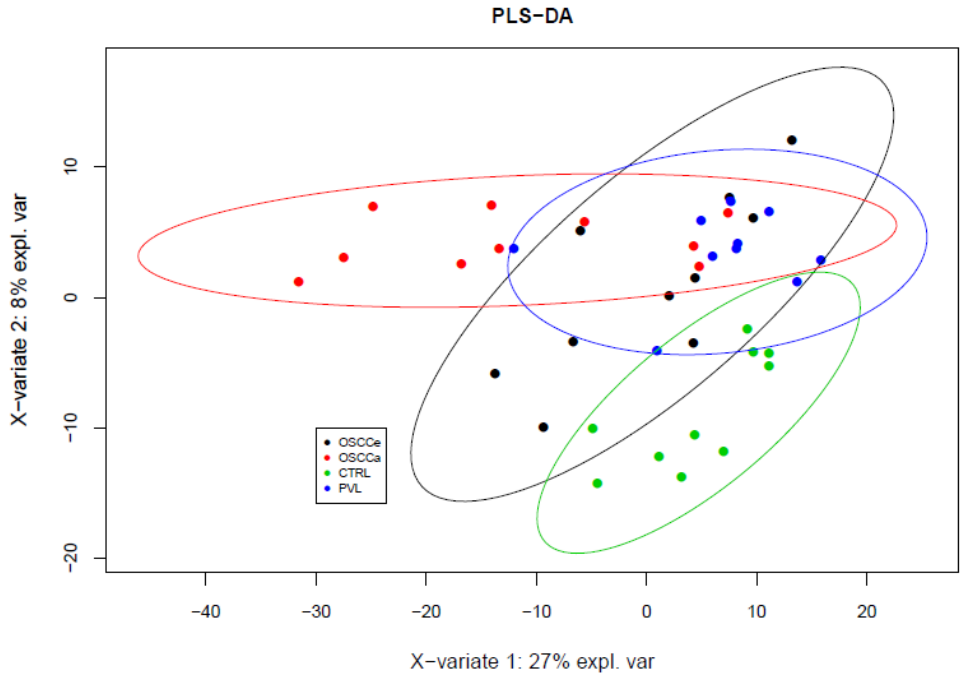


Figure 47. PLS-DA classification plot of whole saliva proteins from controls (green), PVL (blue), early OSCC (black), and advanced OSCC (red) patients based on quantitative SWATH data.

When compared the four groups together, the database contained many more rows than columns in the heat map and PLS-DA did not result in differentiation of PVL and OSCC groups. To reduce the number of the predictive variables (proteins), binomial logistic regression analysis was performed. The examination of the following group pairs included the application of Lasso and Elastic net penalization models, eliminating the non-important proteins according to regression parameters and thus acting as variable selection methods. A cluster analysis based on the Elastic Net (EN) criteria revealed a higher number of proteins than Lasso that could be used to discriminate the studied groups. A heat map was generated for each two by two group comparisons visualizing the resulted proteins expression trends in logarithmically transformed values and samples clustering based on the diagnosis.

3.4.1 *Control vs PVL*

Appling the EN penalization at the database of control and PVL samples resulted in thirteen differential proteins presented on a heat map in Fig.48A. It can be appreciated that the samples are clustered according to their group of belonging. At protein level, two clusters were formed, showing clear differentiation between the two groups. Discriminative analysis classified the samples according to the pertained group, although a certain level of variability is seen within the control group (Fig.48B).

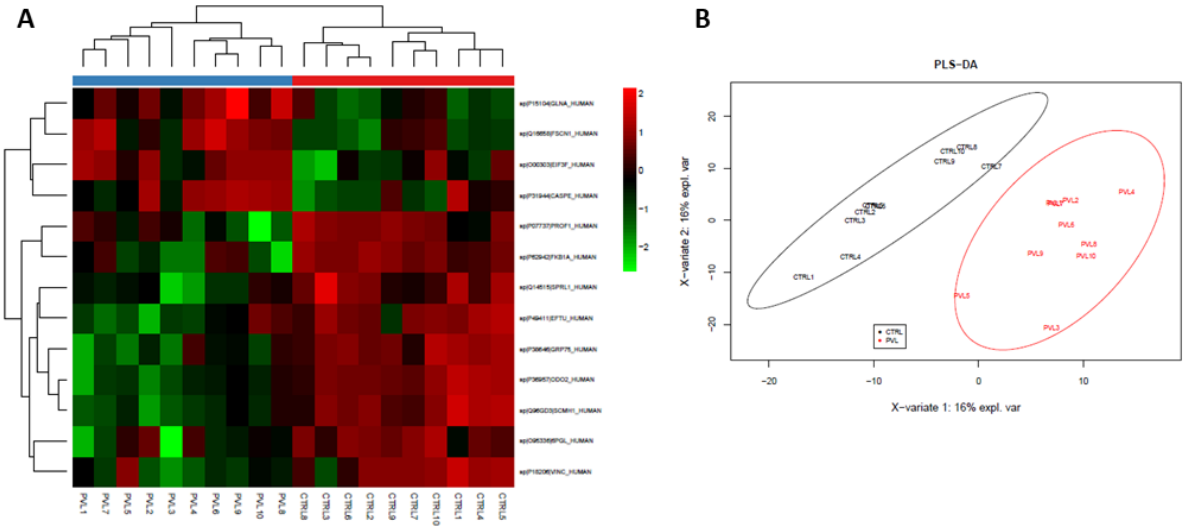


Figure 48. A Heat map and cluster analysis of EN-selected proteins that differ between control (red) and PVL (blue) groups. The heat map scale of Z scores ranges from -2.5 (green) to 2 (red) with a midpoint of 0.0 (black). **B** Classification of control (black) and PVL (red) samples according to PLS-DA.

The summary of the differentially expressed proteins between control and PVL patients is shown in Table 18. Most of them were intracellular proteins, identified with more than two peptide sequences. Intracellular proteins, GLUL, FSCN1, CASP14, and EIF3F were found significantly upregulated (Table 18, dark grey), while nine proteins were downregulated, (Table 18, light grey) in PVL than in control subjects. Sensitivity and specificity analysis described how well the differentially expressed proteins discriminate between controls and PVL patients. Predictive models of these proteins are represented as ROC area under the curves (AUC) (Fig.49, Supplementary figures). Upregulated GLUL, FSCN1, CASP14, EIF3F, as well as downregulated PGLS, SPARKCL1, VCL, HSPA9, TUFM, SCMH1, FKBP1A, DLST, and PFN1 revealed high AUC values (AUC 0.78-0.99)(Table 18).

1Table 18. List of differential proteins between control and PVL groups. The UniProt protein ID and name, are presented. The representative identification with the number (#) of peptides with which a protein was determined, relative abundance (log2Med), FDR adjusted t-test p-value, and ROC AUC value are given in each case.

Protein ID	Protein name	# peptides	Ctrl log2Med	PVL log2Med	*P*-value	FDR	AUC value
P15104	Glutamine synthetase (GLUL)	6	1,09872	2,03313	0,000	0,017	0.96
Q16658	Fascin (FSCN1)	2	-1,34620	-0,54709	0,000	0,003	0.94
P31944	Caspase-14 (CASP14)	2	-2,91580	-1,43862	0,019	0,158	0.79
O00303	Eukaryotic translation initiation factor 3 (EIF3F)	1	-4,54957	-3,32403	0,004	0,074	0.89
O95336	6-phosphogluconolactonase (PGLS)	3	-0,48959	-1,49832	0,001	0,052	0.92
Q14515	SPARC-like protein 1 (SPARKCL1)	3	0,82698	-0,83802	0,000	0,015	0.95
P18206	Vinculin (VCL)	2	-1,22474	-2,50558	0,000	0,034	0.78
P38646	Stress-70 protein (HSPA9)	2	0,34551	-1,02219	0,000	0,006	0.96
P49411	Elongation factor Tu, mitochondrial (TUFM)	1	-1,78584	-3,58591	0,001	0,048	0.91
Q96GD3	Polycomb protein (SCMH1)	1	1,52100	-0,27064	0,000	0,004	0.99
P62942	Peptidyl-prolyl cis-trans isomerase (FKBP1A)	1	-0,76465	-2,42972	0,000	0,026	0.90
P36957	Dihydrolipoyllysine-residue succinyltransferase (DLST)	1	2,37812	0,30745	0,000	0,006	0.99
P07737	Profilin-1 (PFN1)	7	3,19348	2,40510	0,004	0,073	0.88

Log2Med – the average normalized relative abundance of control (Ctrl) and PVL samples (n=10, per group); **FDR** -false discovery rate; **AUC**- area under the curve

3.4.1.1 *Functional analysis*

The nature of differentially expressed proteins between the studied was further characterized by generating protein-protein interaction maps, followed by GO enrichment analysis provided by String bioinformatics resources. The network contained 13 proteins (nodes) and 5 possible functional relations analysed at medium stringency (Fig.50). No biological function was enriched in this loosely connected network. An association to a cellular component (myelin sheath) was suggested between fascin (FSCN1), stress-70 protein (HSPA9), elongation factor Tu (TUFM), dihydro-lipoyl lysine-residue succinyltransferase (DLST), and glutamine synthetase (GLUL).

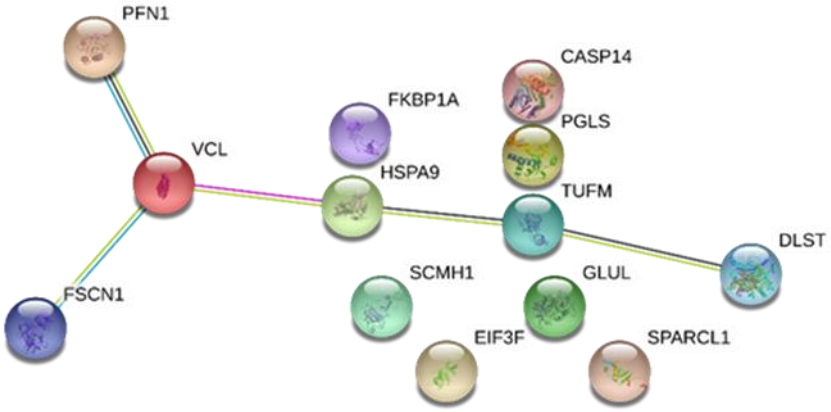


Figure 50. STRING protein-protein interaction network functional enrichment analysis of differentially expressed proteins identified in PVL patients, displayed by String version 11.0 using default settings and medium stringency. Number of nodes/proteins: 13; number of edges: 5; The light blue lines represent database evidence; the purple lines represent experimental evidence, the yellow lines represent text mining evidence, and the black lines represent co-expression evidence

3.4.2 *Control vs early OSCC*

A heat map visualizing differentially expressed proteins between controls and patients at early OSCC is displayed in Fig.51A. It can be seen that samples are clustering per their groups of belonging, except for one control which was classified as early OSCC. EN penalization regression analysis sorted 31 proteins, forming two clumps that distinguish controls from patients at early OSCC. Samples distribution according to the pertaining group is observed as well on a PLS-DA plot (Fig.51B). Information including name, protein ID, and quantitative data of the proteins differentially expressed in early oral cancer stages is summarised in Table 19.

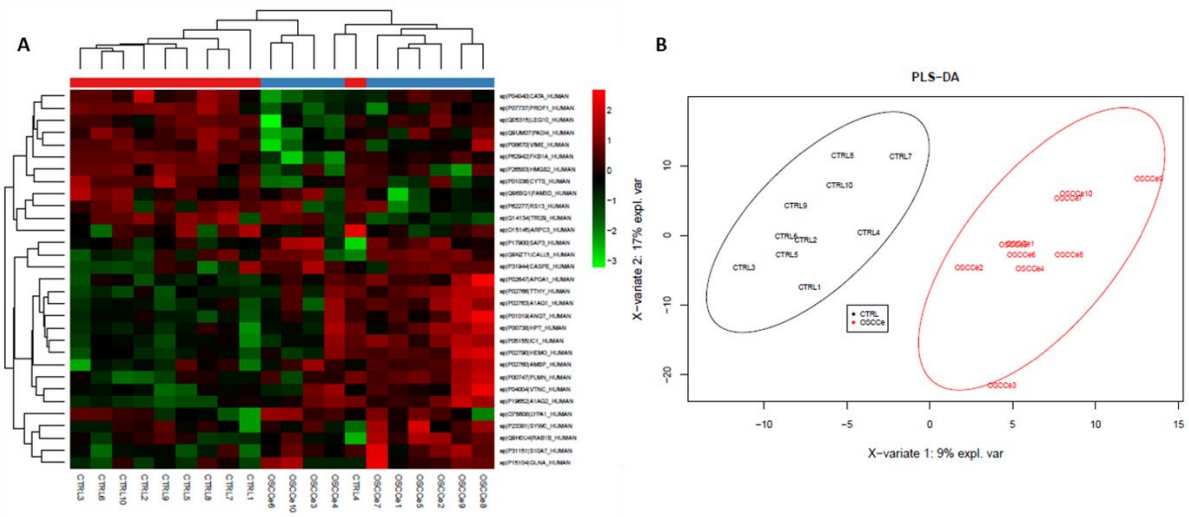


Figure 51. **A** Heat map and cluster analysis of EN-selected proteins discriminating the control (red) and early OSCC (blue) groups. The heat map scale of Z scores ranges from -3 (green) to 2.5 (red) with a midpoint of 0.0 (black). **B** Classification of control (black) and early OSCC (red) samples based on PLS-DA.

From the total of 31 proteins, 18 revealed significantly higher abundance (Table 18, highlighted in dark grey), while 13 proteins were found downregulated (Table 18, highlighted in light grey) among which, 8 were considerably lower in disease compared to healthy controls. ROC analysis were carried out for proteins, differentially expressed in early OSCC stages with p values less than 0.05. The optimal predictive models of 18 upregulated (Fig.52A, Supplementary figures) and 8 downregulated (Fig.52B, Supplementary figures) proteins to discern between health and disease are presented by ROC curves. The candidate markers denoted high AUC values ranging from 0. 79 to 0.91 (Table 19).

Table 19. List of differential proteins between control and early OSCC groups. The UniProt protein ID and name are presented. The representative identification with the number (#) of peptides with which a protein was determined, quantification data (log2Med), FDR adjusted t-test p-value, and ROC AUC value are given in each case.

Protein ID	Protein name	# peptides	Ctrl log2Med	e OSCC log2Med	*P*-value	FDR	AUC value
P00738	Haptoglobin (HP)	25	3,60423	4,87209	0,019	0,740	0.79
P02647	Apolipoprotein A-I (APOA1)	23	3,20770	5,12679	0,003	0,304	0.89
P02790	Hemopexin (HPX)	23	2,91270	4,08118	0,012	0,472	0.83
P31151	Protein S100-A7 (S100A7)	6	0,20957	2,04262	0,001	0,257	0.91
P15104	Glutamine synthetase (GLUL)	6	1,09872	2,05611	0,000	0,017	0.95
P05155	Plasma protease C1 inhibitor (SERPING1)	5	0,74750	1,91991	0,005	0,523	0.83

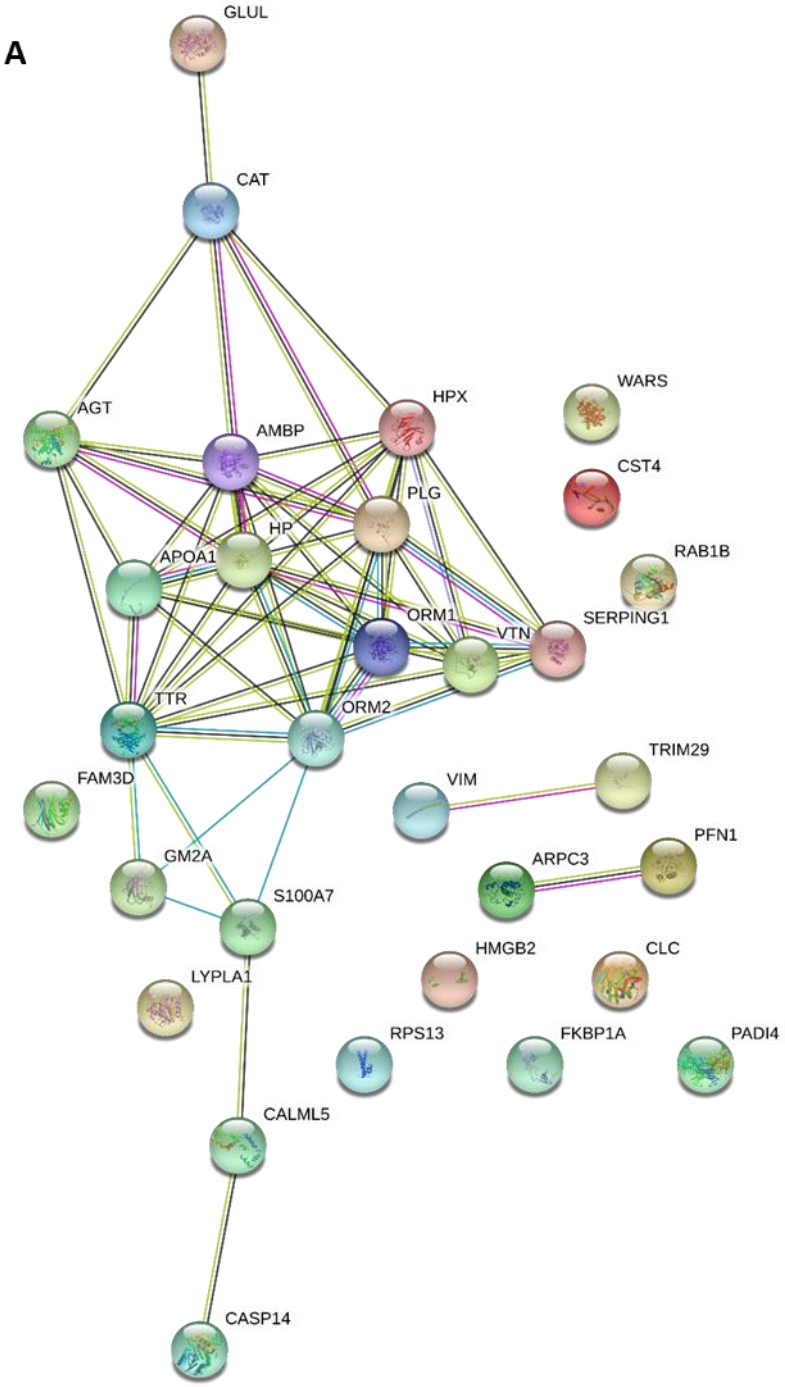
P01019	Angiotensinogen (AGT) (Serpin A8)	4	-1,47282	-0,54400	0,012	0,715	0.87
P02763	Alpha-1-acid glycoprotein 1 (ORM1)	3	1,83116	3,00827	0,005	0,542	0.83
P02766	Transthyretin (TTR)	9	1,63343	2,66232	0,004	0,601	0.84
P04004	Vitronectin (VTN)	3	-0,54775	0,76418	0,019	0,556	0.81
P02760	Protein (AMBP)	3	-0,94109	0,15483	0,002	0,891	0.88
Q9NZT1	Calmodulin-like protein 5 (CALML5)	1	-2,15815	-0,86554	0,003	0,279	0.87
Q9H0U4	Ras-related protein Rab-1B (RAB1B)	2	-3,59031	-2,96119	0,003	0,330	0.86
P31944	Caspase-14 (CASP14)	2	-2,91580	-1,22511	0,008	0,331	0.85
P19652	Alpha-1-acid glycoprotein 2 (ORM2)	2	-0,25768	0,94662	0,003	0,291	0.87
P00747	Plasminogen (PLG)	2	-2,19744	-0,65306	0,014	0,367	0.85
P17900	Ganglioside GM2 activator (GM2A)	2	-0,99859	-0,23336	0,004	0,282	0.88
P23381	Tryptophan--tRNA ligase (WARS)	1	-3,64075	-2,47218	0,020	0,392	0.77
O75608	Acyl-protein thioesterase 1 (LYPLA1)	1	-2,88597	-2,23653	0,064	0,504	-
P08670	Vimentin (VIM)	7	1,82216	0,70968	0,028	0,105	0.86
P07737	Profilin-1 (PFN1)	7	3,19348	2,55143	0,004	0,073	0.88
Q9UM07	Protein-arginine deiminase type-4 (PADI4)	3	0,28529	-0,85725	0,025	0,145	0.80
P04040	Catalase (CAT)	10	2,13757	1,40698	0,022	0,332	0.78
Q14134	Tripartite motif-containing protein 29 (TRIM29)	4	1,22068	-0,00939	0,018	0,768	0.78
Q05315	Galectin-10 (CLC)	3	0,040364	-0,844096	0,068	0,431	-
P01036	Cystatin-S (CST4)	2	3,67884	0,48451	0,002	0,477	0.89
P26583	High mobility group protein B2 (HMGB2)	2	0,08814	-1,49674	0,011	0,359	0.70
P62942	Peptidyl-prolyl cis-trans isomerase (FKBP1A)	1	-0,76465	-2,19694	0,008	0,334	0.84
P62277	40S ribosomal protein S13 (RPS13)	2	-0,89755	-1,59949	0,069	0,499	-
Q96BQ1	Protein (FAM3D)	2	-0,45627	-0,98242	0,069	0,503	-
O15145	Actin-related protein 3 (ARPC3)	1	-1,89475	-3,14059	0,259	0,760	-

Log2Med – average normalized relative abundance of control (Ctrl) and early (e) OSCC samples (n=10 per group); **FDR**- false discovery rate; **AUC**- area under the curve

3.4.2.1 *Functional analysis*

For a more detailed functional analysis of the differentially expressed proteins, gene ontology (GO) analysis was performed; the biological process and molecular function according to GO (http://www.geneontology.org/) were examined. The altered 31 proteins showed a highly interconnected protein-protein interaction network with 58 interactions analysed at medium stringency (Fig.53A). As summarized in Fig.53B, the mapped genes were associated with 124 biological processes and 11 molecular functions. The highest rank biologic processes were vesicle-mediated transport and secretion (exocytosis) (FDR < 0.005) and immune response. The highest association with the molecular function was

inhibition of enzymatic activities, and receptor-associated protein activity (signaling) (FDR < 0.05).



B

GO term ID	Pathway description	Observed gene count	FDR
I.Biological process			
GOTERM_BP first 9 out of 124 chart records			
GO:0016192	vesicle-mediated transport	16	1.56e-06
GO:0045055	regulated exocytosis	11	4.25e-06
GO:0032940	secretion by cell	12	6.74e-06
GO:0006955	immune response	14	1.06e-05
GO:0002376	immune system process	16	2.63e-05
GO:0006810	transport	20	4.81e-05
GO:0002443	leukocyte mediated immunity	9	7.19e-05
GO:0002252	immune effector process	10	0.00010
GO:0002682	regulation of immune system process	12	0.00010
II. Molecular function			
GOTERM_MF first 6 out of 11 chart records			
GO:0004857	enzyme inhibitor activity	6	0.0077
GO:0004866	endopeptidase inhibitor activity	4	0.0077
GO:0005102	signaling receptor binding	10	0.0077
GO:0005515	protein binding	21	0.0077
GO:0030234	enzyme regulator activity	8	0.0077
GO:0042802	identical protein binding	11	0.0077

Figure 53. The protein-protein interaction network and functional classification (**A**) of up and down-regulated proteins in early OSCC. Number of nodes: 31; number of edges: 58; each node represents a protein and the edges represent the interconnectivity. The enrichment table of GO terms (**B**) calculated by String of differentially expressed proteins is shown indicating the number of proteins belonging to each term and the false discovery rate (FDR).

3.4.3 *Control vs advanced OSCC*

Binominal logistic regression analysis revealed 29 proteins with the potential to differ patients at advanced OSCC stages from their healthy counterparts. On the heat map shown in Fig.54A can be seen that the samples were well grouped as per control and advanced OSCC and a series of differential proteins with discriminating profiles between the investigated groups. Case-control sample arrangement was further noted on a PLS-DA classification plot (Fig.54B).

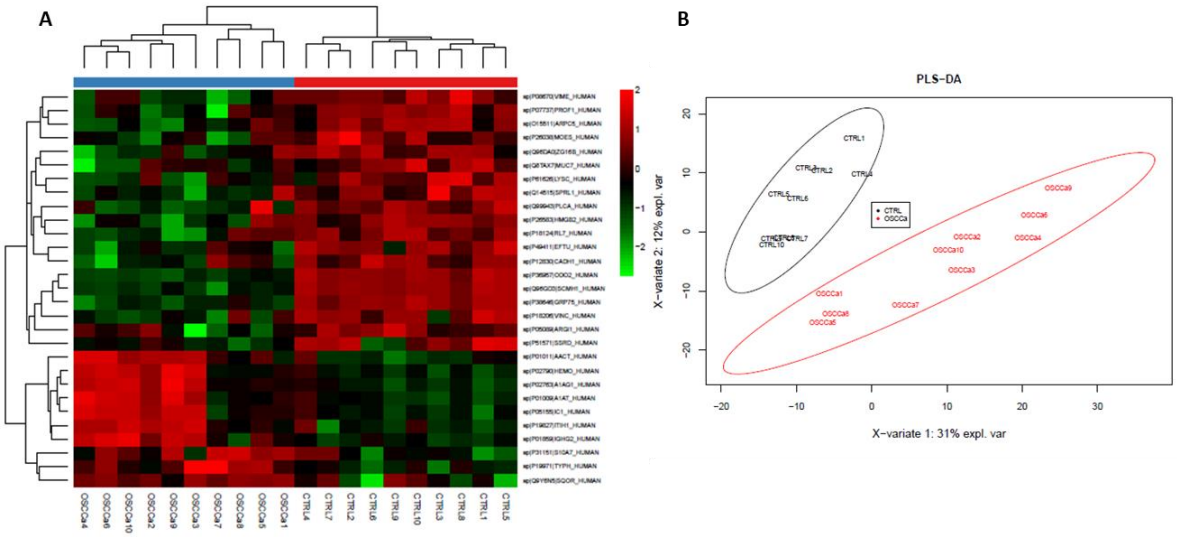


Figure 54. A Heat map and cluster analysis showing EN-selected proteins distinctive for control and advanced OSCC groups. Clusters of control (red) and advanced OSCC (blue) samples are shown at the top of the heat map. The heat map scale of Z scores ranges from -2.5 (green) to 2 (red) with a midpoint of 0.0 (black). **B** PLS-DA classification plot of proteins from control (black) and advanced OSCC (red) samples.

A summary of the 29 proteins with altered abundance in patients at advanced OSCC group is displayed in Table 20. The comparative analysis determined 10 representatives with increased abundance (Table 20, dark grey), whereas, 16 out of the 29 proteins with reduced expression (Table 20, light grey) were considerably low in disease. ROC analysis were used to determine the discriminatory efficacy of statistically significant, differentially expressed proteins between control and advanced OSCC individuals. The diagnostic performance of 10 up- (Fig.55A, Supplementary figures) and 16 downregulated (Fig.55B, Supplementary figures) markers in advanced OSCC is represented by AUC values reaching up to 1.00 (Table 19).

Table 20. List of differential proteins between control and advanced OSCC groups. The UniProt protein ID and name are presented. The representative identification with the number (#) of peptides with which a protein was determined, relative abundance (log2Med), FDR adjusted t-test p-value, and ROC AUC value are given in each case. **Log2Med** – the average normalized relative abundance of controls (Ctrl) and advanced (adv OSCC) samples (n=10 per group); **FDR**- false discovery rate; **AUC**- area under the curve.

Protein ID	Protein name	# peptides	Ctrl log2Med	adv OSCC log2Med	*P*-value	FDR	AUC value
P01009	Alpha-1-antitrypsin (SERPINA1)	27	4,86531	6,90271	0,002	0,071	0.88
P02790	Hemopexin (HPX)	14	2,91270	5,10540	0,002	0,074	0.91
P01011	Alpha-1-antichymotrypsin (GIG25)	9	1,74022	3,11680	0,002	0,071	0.87

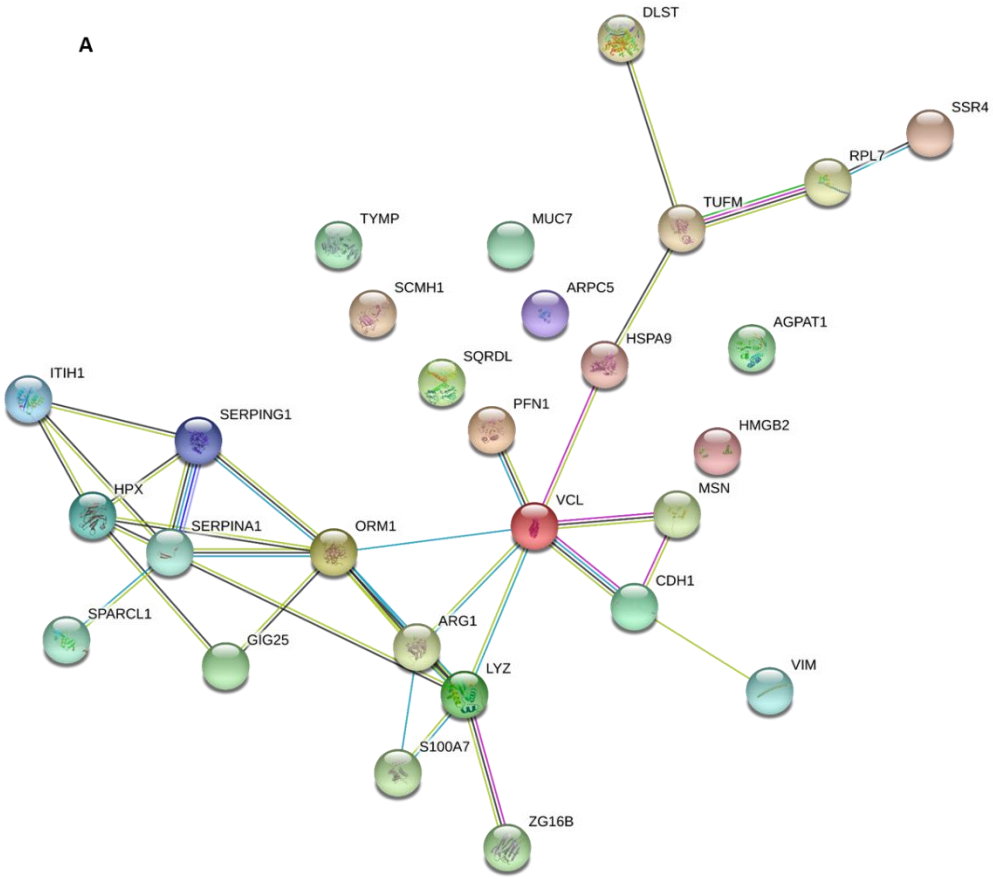
P01859	Immunoglobulin heavy constant gamma 2 (IGHG2)	6	4,72908	6,58322	0,002	0,072	0.88
P31151	Protein S100-A7 (S100A7)	6	0,20957	2,00364	0,001	0,081	0.87
Q9Y6N5	Sulfide:quinone oxidoreductase (SQRDL)	3	-1,18584	-0,08147	0,001	0,083	0.97
P02763	Alpha-1-acid glycoprotein 1 (ORM1)	3	1,83116	4,10635	0,000	0,052	0.96
P05155	Plasma protease C1 inhibitor (SERPING1)	5	0,74750	2,82388	0,002	0,075	0.89
P19971	Thymidine phosphorylase (TYMP)	4	-0,83235	0,46389	0,000	0,008	0.96
P19827	Inter-alpha-trypsin inhibitor heavy chain H1 (ITIH1)	2	-2,30813	-0,23171	0,001	0,076	0.91
P61626	Lysozyme C (LYZ)	8	6,63884	5,46381	0,005	0,098	0.84
P26038	Moesin (MSN)	9	1,16920	0,88757	0,186	0,486	-
Q96DA0	Zymogen granule protein 16 homolog B (ZG16B)	10	9,86591	8,51886	0,002	0,072	0.90
P08670	Vimentin (VIM)	7	1,82216	1,09198	0,012	0,132	0.83
P07737	Profilin-1 (PFN1)	7	3,19348	2,57842	0,011	0,134	0.86
Q8TAX7	Mucin-7 (MUC7)	2	0,98353	-0,33505	0,012	0,133	0.82
P18206	Vinculin (VCL)	2	-1,22474	-2,26852	0,005	0,095	0.79
P18124	60S ribosomal protein L7 (RPL7)	2	0,89230	0,03433	0,000	0,025	0.97
P05089	Arginase-1 (ARG1)	2	-1,89678	-2,47578	0,111	0,384	-
P49411	Elongation factor Tu (TUFM)	1	-1,78584	-3,46530	0,000	0,046	0.90
P26583	High mobility group protein B2 (HMGB2)	2	0,08814	-1,52601	0,001	0,076	0.82
P51571	Translocon-associated protein subunit delta (SSR4)	1	-3,88640	-5,05475	0,051	0,262	-
P12830	Cadherin-1 (CDH1)	2	-0,33364	-0,80704	0,021	0,168	0.70
Q14515	SPARC-like protein 1 (SPARCL1)	3	0,82698	-0,33638	0,007	0,111	0.81
O15511	Actin-related protein 2/3 complex subunit 5 (ARPC5)	1	0,85735	0,17522	0,000	0,007	0.96
P38646	Stress-70 protein (HSPA9)	2	0,34551	-2,22214	0,000	0,000	1.00
Q96GD3	Polycomb protein (SCMH1)	1	1,52100	-2,84126	0,000	0,000	1.00
P36957	Dihydrolipoyllysine-residue succinyltransferase (DLST)	1	2,37812	-0,94362	0,000	0,000	1.00
Q99943	Protein G15 (AGPAT1)	1	3,03251	0,75315	0,001	0,071	0.87

Log2Med – average normalized abundance of controls (Ctrl) and advanced (adv OSCC) samples (n=10 per group); **FDR**- false discovery rate.

3.4.3.1 *Functional analysis*

The String network analysis of 29 proteins with altered abundance demonstrated 32 possible interactions (Fig.356A). A cluster of proteins involved in the regulation of exocytosis including SERPING1, SERPINA1, GIG25, ORM1, ARG, LYY, S100A7, and VCL was observed. The highest rank biologic processes and association with molecular functions (both, FDR < 0.05) are reported in Fig. 34B. The enriched functions indicate active regulatory mechanisms implicating the immune system, inhibition of endopeptidase enzymatic activity as well structural constituents of the actin cytoskeleton (Fig.56B).

A



B

GO term ID	Pathway description	Observed gene count	FDR
I. Biological process (1/ 29 genes not included)			
GOTERM_BP first 7 out of 22 chart records			
GO:0002376	immune system process	14	0.00021
GO:0002443	leukocyte mediated immunity	9	0.00021
GO:0006955	immune response	12	0.00021
GO:0043312	neutrophil degranulation	8	0.00021
GO:0045055	regulated exocytosis	9	0.00021
GO:0045321	leukocyte activation	9	0.00021
GO:0002682	regulation of immune system process	10	0.00075
II. Molecular function			
COTERM_MF 5 charts total			
GO:0004867	serine-type endopeptidase inhibitor activity	4	0.0017
GO:0050786	RAGE receptor binding	2	0.0080
GO:0005200	structural constituent of cytoskeleton	3	0.0111
GO:0003779	actin binding	4	0.0449
GO:0005198	structural molecule activity	5	0.0449

Figure 56. A Functional interaction network of 28 identified proteins with differential abundance in advanced OSCC showing 32 interconnections (edges) and **B** the enrichment table of GO terms indicating the number of the proteins belonging to each term and the FDR calculated by String annotation tool.

3.4.4 *PVL vs early OSCC*

EN penalized binary regression analysis of control and PVL database resulted in a group of twelve differential proteins shown via a heat map on Fig.57A. It can be perceived that the samples were assembled in compliance with the related pathology groups. At protein level, differential expression was observed in two clusters. Discriminative analysis classified the samples as PVL and early OSCC group (Fig.57B), though, the last disclosed relatively high variability between the individual samples.

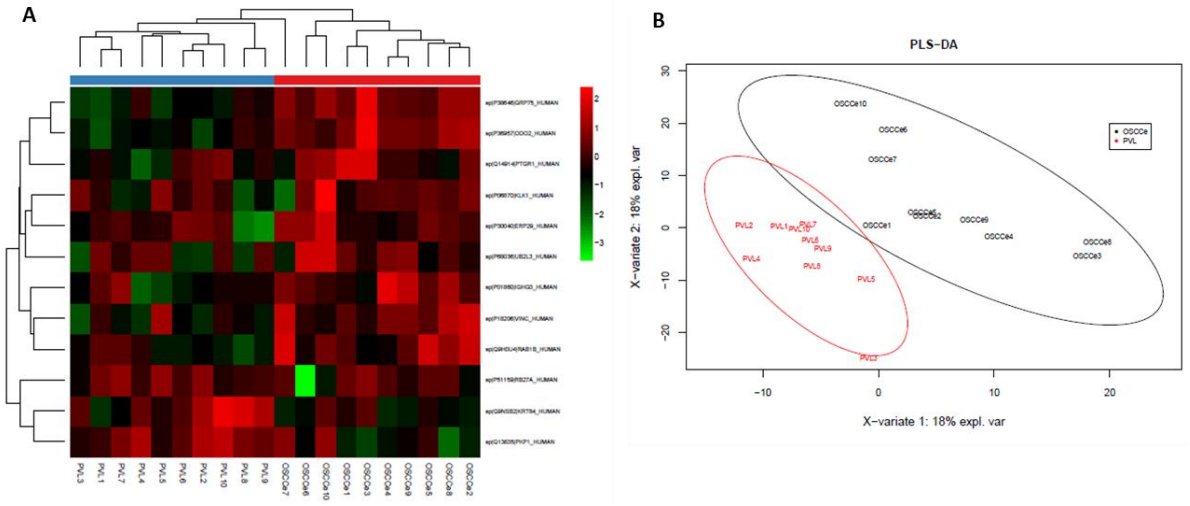


Figure 57. A Heat map and cluster analysis showing EN-selected proteins distinctive for PVL and early OSCC groups. Clusters of PVL (red) and early OSCC (blue) samples are shown at the top of the heat map. The heat map scale of Z scores ranges from -3.5 (green) to 2.5 (red) with a midpoint of 0.0 (black). **B** Classification plot of PVL (red) and early OSCC (black) samples according to PLS-DA.

Among the differential proteins, 9 exhibited elevated (Table 21, dark grey) and 3 declined expression (Table 21, light grey) in early OSCC than in patients with PVL. Ras-related protein Rab-1B (RAB1B), vinculin (VCL), prostaglandin reductase 1 (PTGR1), immunoglobulin heavy constant gamma (IGHG), stress-70 protein (HSPA9), and Dihydrolipoamide S-succinyltransferase (DLST) were observed significantly upregulated, whereas keratin, type II (KRT84), and plakophilin-1 (PKL1) showed considerable downregulation. Sensitivity and specificity estimation demonstrated the capacity of 6 proteins with significantly higher

(Fig.58A, Supplementary figures) and 2 with significantly lower (Fig58B, Supplementary figures) expressions to discriminate between patients with PVL and early stages OSCC lesions. ROC curve analysis revealed valuable candidates demonstrating AUC values greater than 0.7.

Table 21. List of differential proteins between PVL and early OSCC groups. The UniProt protein ID and name are presented. The representative identification with the number (#) of peptides with which a protein was determined, quantification data (log2Med), FDR adjusted t-test p-value, and ROC AUC value are given in each case.

Protein ID	Protein name	# peptides	PVL log2Med	e OSCC log2Med	p-value	FDR	AUC value
P06870	Kallikrein-1 (KLK1)	9	2,40589	3,02643	0,215	0,809	-
Q9H0U4	Ras-related protein Rab-1B (RAB1B)	2	-3,51928	-2,96119	0,002	0,343	0.89
P18206	Vinculin (VCL)	2	-2,50558	-1,48434	0,002	0,299	0.88
Q14914	Prostaglandin reductase 1 (PTGR1)	2	-1,38119	-0,42372	0,047	0,667	0.75
P01860	Immunoglobulin heavy constant gamma 3 (IGHG3)	4	0,40378	1,73537	0,008	0,512	0.87
P68036	Ubiquitin-conjugating enzyme E2 (UBE2L3)	1	-4,40966	-3,50060	0,072	0,700	-
P38646	Stress-70 protein (HSPA9)	2	-1,02219	0,78660	0,000	0,014	1.00
P30040	Endoplasmic reticulum resident protein 29 (ERP29)	1	-2,40397	-1,81484	0,054	0,668	-
P36957	Dihydrolipoamide S-succinyltransferase (DLST)	1	0,30745	2,47004	0,000	0,029	0.99
Q9NSB2	Keratin, type II cuticular Hb4 (KRT84)	12	2,47599	0,77594	0,018	0,634	0.72
Q13835	Plakophilin-1 (PKP1)	6	2,58738	1,69214	0,001	0,254	0.80
P51159	Ras-related protein Rab-27A (RAB27A)	2	0,98253	0,47872	0,279	0,793	-

Log2Med – average normalized relative abundance of PVL and early (e) OSCC samples (n=10 per group); **FDR**-false discovery rate; **AUC**- area under the curve

3.4.4.1 *Functional analysis*

The network of PVL differentially expressed proteins contained 11 components (nodes) with only 3 possible relations indicating that there were no significant interactions (Fig.59). No enrichment of biological function was indicated, either. However, a functional connection was suggested between HSPA9 - mitochondrial chaperone protein playing an important role in the mitochondrial iron-sulfur cluster (ISC) biogenesis; VCL - actin filament (F-actin)-a binding protein involved in cell-matrix adhesion and cell-cell adhesion; RAB27A - plays a role in cytotoxic granule exocytosis in lymphocytes, and RAB1B - a small protein with GTPase activity, key regulators of intracellular membrane trafficking.

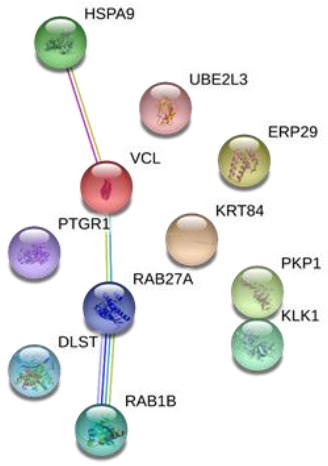


Figure 59. Functional interaction network of differentially expressed proteins in PVL generated by String 11.0 using default settings and medium stringency. Number of nodes: 11; number of edges: 3;

3.4.5 *PVL vs advanced OSCC*

EN regularization of protein expression data obtained from advanced OSCC and PVL samples was represented with a heat map on Fig.60A, revealing ten proteins sorted in two clusters with differential expression patterns within the two groups. PLS-DA affirmed sample grouping after the pathology group of belonging (Fig.60B).

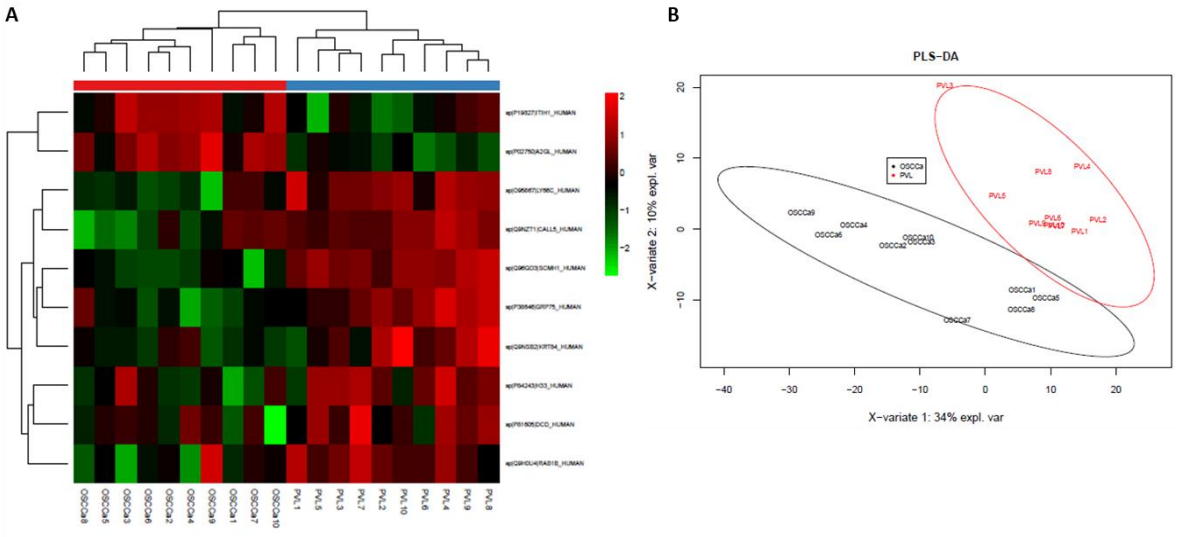


Figure 60. **A** Heat map and cluster analysis of EN-selected proteins distinctive for the PVL (blue) and advanced OSCC (red) groups are shown at the top of the heat map. The heat map scale of Z scores ranges from -2.5 (green) to 2.5 (red) with a midpoint of 0.0 (black). **B** PLS-DA classification of PVL (red) and early OSCC (black) samples.

Amongst the outlined proteins, 2 were notably increased in advanced OSCC (Table 22, dark grey). Significantly reduced abundance was estimated in 3 out of the 7 downregulated proteins (Table22, light grey). Gene ontology analysis did not reveal any functional relation

among the 10 mapped genes. ROC curves represent predictive models of 2 significantly increased (Fig.61A, Supplementary figures) and 3 decreased (Fig.61B, Supplementary figures) proteins in advanced OSCC compared to PVL patients. The markers exhibited discriminatory capacity reported by AUC values ranging from 0.83 to 0.99 (Table 22).

Table 22. List of differential proteins between PVL and advanced OSCC groups. The UniProt protein ID and name are presented. The representative identification with the number (#) of peptides with which a protein was determined, relative abundance (log2Med), FDR adjusted t-test *p*-value, and ROC AUC value are given in each case.

Protein ID	Protein name	# peptides	PVL log2Med	adv OSCC log2Med	*P*-value	FDR	AUC value
P02750	Leucine-rich alpha-2-glycoprotein (LRG1)	2	-1,28119	1,13842	0,000	0,016	0.94
P19827	Inter-alpha-trypsin inhibitor heavy chain H1 (ITIH1)	2	-2,20847	-0,23171	0,003	0,088	0.83
Q9NSB2	Keratin, type II cuticular (KRT84)	12	2,47599	0,85392	0,171	0,425	-
Q9H0U4	Ras-related protein Rab-1B (RAB1B)	2	-3,51928	-3,84041	0,261	0,501	-
P81605	Dermcidin (DCD)	1	-2,93484	-3,64133	0,183	0,433	-
Q9NZT1	Calmodulin-like protein 5 (CALML5)	6	-0,37651	-3,16167	0,240	0,480	-
O95867	Lymphocyte antigen 6 complex (LY6G6C)	2	1,01975	-0,93070	0,000	0,007	0.87
P38646	Stress-70 protein (HSPA9)	2	-1,02219	-2,22214	0,000	0,023	0.93
P84243	Histone H3.3 (H3F3B)	1	-3,89741	-4,77885	0,097	0,314	-
Q96GD3	Polycomb protein (SCMH1)	1	-0,27064	-2,84126	0,000	0,005	0.99

Log2Med – average normalized relative abundance for PVL and advanced (adv) OSCC samples (n=10 per group); **FDR**- false discovery rate; **AUC**- area under the curve

3.4.6 *PVL vs OSCC*

A comparison of patients OSCC (at initial and advanced stages) versus PVL ones, EN penalty regression resulted in a series of differential proteins, the expression trends of which were depicted with a matrix, displayed in Fig.62A. It was observed that samples are gathered as per the pertained group, yet, three of them seemed more similar to PVL than OSCC, according to the protein presentation levels. The resulted proteins can differ between the two pathology groups. Furthermore, different protein patterns were noted within the OSCC cohort, between the early and advanced stage samples. PLS-DA demonstrated sample classification according to the pathology of belonging, however, both groups are not exclusive (Fig.62B).

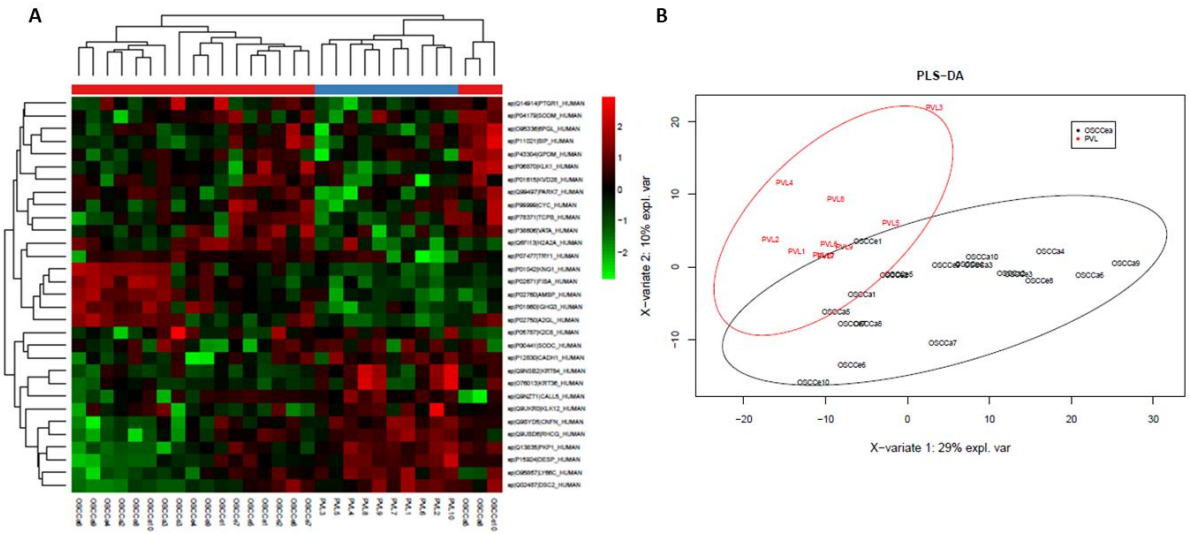


Figure 62. A Heat map depicting differential proteins after EN penalty that differ between OSCC (red) and PVL (blue) samples, clustered on the top of the heat map. The heat map scale of Z scores ranges from -2.5 (green) to 2.5 (red) with a midpoint of 0.0 (black). **B** Classification of OSCC (n=20) (black) and PVL (n=10) (red) samples according to PLS-DA.

Statistical analysis revealed 18 proteins with increased (Table 23, dark grey) and 13 with decreased (Table 23, light grey) abundance in OSCC than in PVL cases. Among the significantly overexpressed compounds were fibrinogen alpha chain (FGA), kallikrein-1 (KLK1), immunoglobulin heavy constant gamma 3 (IGHG3), leucine-rich alpha-2-glycoprotein LRG1), etc. Notably lower expression was estimated in desmoplakin (DSP), plakophilin-1 (PKP1), desmocollin-2 (DSC2) cornifelin (CNFN), and others. Prediction model for sensitivity and specificity performance of proteins with a significantly different expression between PVL and OSCC are represented with ROC curves. 13 OSCC upregulated proteins (Fig.63A, Supplementary figures) displayed AUCs in the range from 0.75 to 0.83 and 6 OSCC downregulated proteins (Fig.63B, Supplementary figures) from 0.61 to 0.85 suggesting a relatively high capacity of the studied marked to discern between PVL and OSCC patients. Four upregulated proteins including PGLS, IGHG3, LRG1, and AMBP noted 0.75, 0.86, 0.86, and 0.82 AUCs, respectively, and 4 downregulated proteins including KRT84, KLK, CALML5, and LY6G6C marked 0.71, 0.61, 0.62, and 0.79, respectively (Table 23). However, some proteins exhibited higher predictive performance within the aforementioned group comparisons, therefore, ROC graphs are omitted to avoid redundancy.

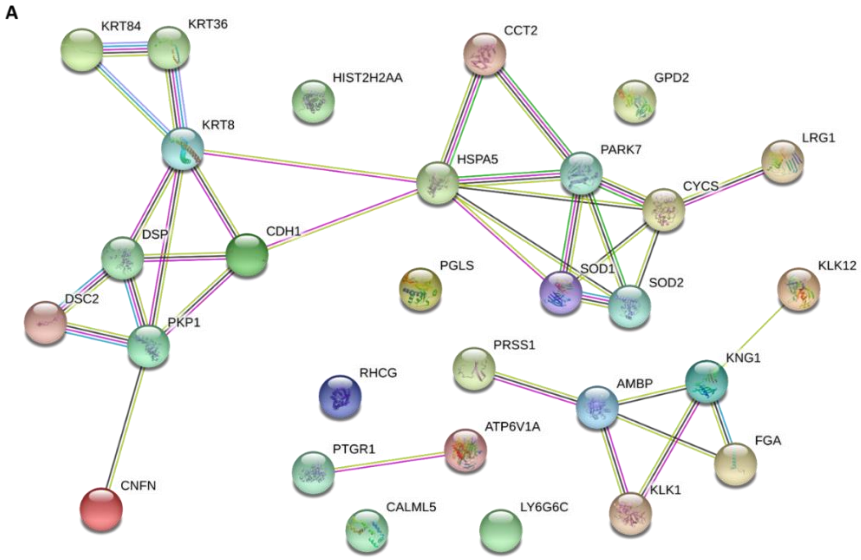
Table 23. List of differential proteins between PVL and OSCC groups. The UniProt protein ID and name are presented. The representative identification with the number (#) of peptides with which a protein was determined, quantification data (log2Med), FDR adjusted t-test p-value, and ROC AUC value are given in each case.

Protein ID	Protein name	# peptides	PVL log2Med	OSCC log2Med	*P*-value	FDR	AUC value
P11021	Endoplasmic reticulum chaperone BiP (HSPA5)	13	1,88614	2,32352	0,003	0,097	0.79
P02671	Fibrinogen alpha chain (FGA)	7	0,79543	2,40911	0,001	0,035	0.83
P06870	Kallikrein-1 (KLK1)	9	2,40589	3,27298	0,041	0,232	0.75
Q14914	Prostaglandin reductase 1 (PTGR1)	2	-1,38119	-0,51768	0,016	0,199	0.77
P04179	Superoxide dismutase (SOD2)	3	-0,47552	0,02754	0,062	0,283	-
P78371	T-complex protein 1 subunit beta (CCT2)	2	-3,98182	-3,28098	0,013	0,206	0.78
O95336	6-phosphogluconolactonase (PGLS)	3	-1,49832	-0,84911	0,028	0,222	0.75
P01860	Immunoglobulin heavy constant gamma 3 (IGHG3)	4	0,40378	2,12993	0,001	0,036	0.87
P38606	V-type proton ATPase catalytic subunit A (ATP6V1A)	1	-3,36350	-2,43967	0,021	0,202	0.78
P02750	Leucine-rich alpha-2-glycoprotein (LRG1)	2	-1,28119	0,44406	0,000	0,021	0.94
P99999	Cytochrome c (CYCS)	1	-4,15015	-3,50759	0,089	0,341	-
P01042	Kininogen-1 (KNG1)	3	-2,60074	-1,16315	0,002	0,070	0.79
P02760	Protein (AMBP)	3	-0,80313	0,43453	0,004	0,109	0.82
Q99497	Protein/nucleic acid deglycase DJ-1(PARK7)	1	-2,84376	-1,70670	0,015	0,208	0.82
P01615	Immunoglobulin kappa variable 2D-28 (IGKV2D)	2	0,35255	1,14823	0,033	0,220	0.78
P07477	Trypsin-1 (PRSS1)	1	1,68938	4,25564	0,074	0,324	-
P43304	Glycerol-3-phosphate dehydrogenase (GPD2)	1	-3,70382	-2,99409	0,042	0,232	0.77
Q6FI13	Histone H2A type 2-A (HIST2H2AA)	1	-4,40336	-3,03841	0,015	0,209	0.80
P15924	Desmoplakin (DSP)	29	4,54379	3,52273	0,000	0,019	0.85
Q9NSB2	Keratin, type II cuticular (KRT84)	12	2,47599	0,81493	0,017	0,195	0.72
O76013	Keratin, type I cuticular (KRT36)	5	1,04905	-0,28483	0,054	0,255	-
Q13835	Plakophilin-1 (PKP1)	6	2,58738	1,42242	0,000	0,002	0.83
Q02487	Desmocollin-2 (DSC2)	5	1,69905	1,14683	0,000	0,029	0.74
Q9UBD6	Ammonium transporter Rh type C (RHCG)	4	1,82689	0,81554	0,000	0,003	0.85
P00441	Superoxide dismutase (SOD1)	1	-3,30847	-3,65845	0,412	0,695	-
Q9UKR0	Kallikrein-12 (KLK12)	2	-0,42817	-1,27642	0,043	0,232	0.61
P05787	Keratin, type II cytoskeletal 8 (KRT8)	1	-3,17979	-3,46608	0,483	0,738	-
P12830	Cadherin-1 (CDH1)	2	-0,64351	-0,90080	0,173	0,450	-
Q9NZT1	Calmodulin-like protein 5 (CALML5)	1	-0,37651	-2,01361	0,007	0,143	0.74
Q9BYD5	Cornifelin (CNFN)	2	0,21766	-0,59762	0,000	0,018	0.82
O95867	Lymphocyte antigen 6 complex locus protein G6c (LY6G6C)	2	1,01975	-0,38584	0,000	0,021	0.87

Log2Med – average normalized abundance of PVL and OSCC samples (n= 10 and 20, respectively); **FDR**-false discovery rate; **AUC** – area under the curve

3.4.6.1 *Functional analysis*

Proteins with altered expression in patients with OSCC at different clinical stages compared to the ones with premalignant (PVL) lesions were subjected to gene ontology screening. String functional protein association network identified 29 proteins and 35 possible interconnections (Fig.64A). GO enrichment analysis of the mapped genes revealed associations with 85 biological processes and 11 molecular functions, the highest ranks of which (FDR < 0.05) are summarized in Fig.64B. Enrichment for biological process revealed epidermis development through epithelial cell differentiations including members of the keratin family (KRT 84, KRT 36, and KRT 8), desmosome components (DSP and DSC2), and components of cell junctions and stratified squamous epithelia (PKP1 and CNFN). Other enriched processes involve programmed cell death represented by proteins CYCS, SOD2, HSPA5, KLK12, KRT 84, KRT 36, KRT 8, DSP, DSC2, and PKP1 and regulation of exocytosis implicating DSP, PKP1, CCT2, SOD1, LRG1, KNG1, FGA and CALML5. Molecular function enrichment disclosed structural constituent of skin epidermis (KRT 84, KRT 36, and PKP1), antioxidant and oxidoreductase activity (SOD1, SOD2, PARK, CYCS, and GDP2).



B

Pathway term ID	Pathway description	Observed gene count	FDR
I.Biological process			
GOTERM_BP first 8 out of 85 chart records			
GO:0008544	epidermis development	10	3.21e-07
GO:0009913	epidermal cell differentiation	9	3.21e-07
GO:0070268	cornification	7	3.21e-07
GO:0031424	keratinization	8	4.53e-07
GO:0030855	epithelial cell differentiation	10	4.78e-06
GO:0012501	programmed cell death	10	0.00028
GO:0045055	regulated exocytosis	8	0.00075
GO:0032940	secretion by cell	9	0.00078
II. Molecular function			
GOTERM_MF 5 out of 11 chart records			
GO:0030280	structural constituent of epidermis	3	0.00035
GO:0004784	superoxide dismutase activity	2	0.0040
GO:0097110	scaffold protein binding	3	0.0040
GO:0016209	antioxidant activity	3	0.0061
GO:0016491	oxidoreductase activity	6	0.0092

Figure 64. Protein-protein interaction network and functional classification (**A**) of up and down-regulated proteins in OSCC. Number of identified nodes: 29 out of 31; number of edges: 35; The enrichment table of GO terms (**B**) is shown indicating the number of proteins belonging to each term and the false discovery rate (FDR) recognized by the STRING annotation tool.

3.4.7 *Early vs advanced OSCC*

A heat map illustrating selected proteins after EN-penalized regression with the potential to discriminate patients at early from advanced OSCC stages is shown in Fig.65A. According to the varying protein levels, samples were clustered correspondingly to the initial and advanced OSCC stages, appreciated also in the PLS-DA classification model (Fig.65B).

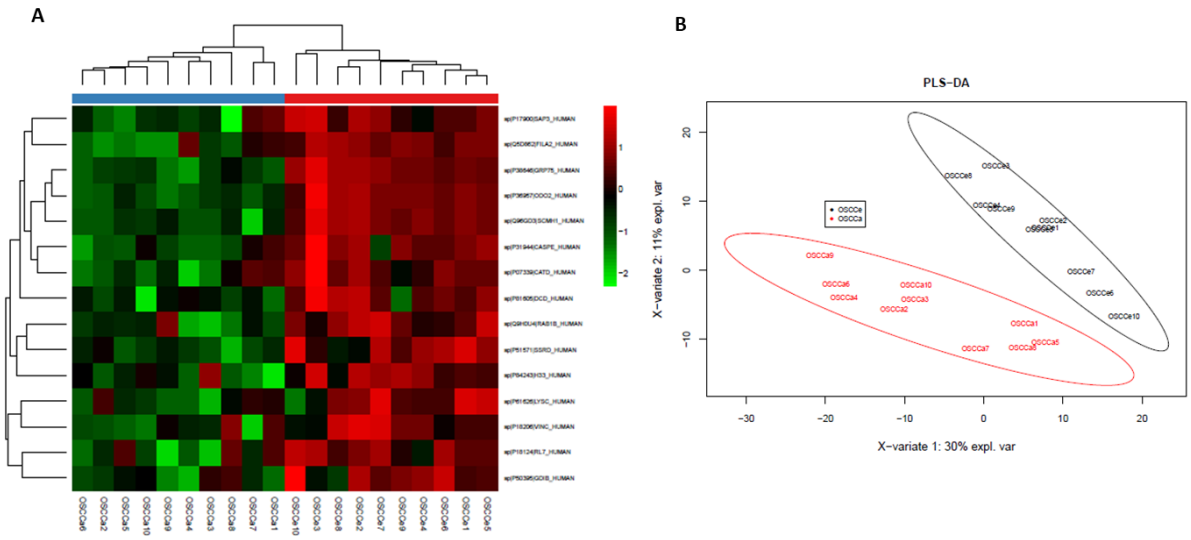


Figure 65. A Heat map of differentially expressed proteins after EN penalty in a comparison between early OSCC (red) and advanced OSCC (blue) samples, clustered on the top of the heat map. The heat map scale of Z scores ranges from -2 (green) to 2 (red) with a midpoint of 0.0 (black). **B** Classification of early OSCC (black) and advanced OSCC (red) samples based on protein expression PLS-DA.

The summary of differentially expressed protein IDs, names, subcellular location, and quantitative data is represented in Table 24. It can be appreciated that all, except Rab GDP dissociation inhibitor beta (GDI2), exhibited significantly increased abundance within the early OSCC compared to advance OSCC stages with *P* values less than 0.05. ROC analysis were carried out to determine the discriminatory efficacy of statistically significant, differentially expressed proteins between early and advanced OSCC groups. The diagnostic performance of 11 early OSCC upregulated markers disclosed AUC values higher than 0.7 (Fig.66, Supplementary figures).

Table 24. List of differential proteins between early and advanced OSCC groups. The UniProt protein ID and name are presented. The representative identification with the number (#) of peptides with which a protein was determined, relative abundance, FDR adjusted t-test *p*-value, and ROC AUC value are given in each case.

Protein ID	Protein name	# peptides	e OSCC log2Med	adv OSCC log2Med	*P*-value	FDR	AUC value
P50395	Rab GDP dissociation inhibitor beta (GDI2)	6	1,45987	1,24021	0,114	0,874	-
P07339	Cathepsin D (CTSD)	3	2,27887	1,40969	0,002	0,201	0.77
Q9H0U4	Ras-related protein Rab-1B (RAB1B)	2	-2,96119	-3,84041	0,007	0,402	0.87
P81605	Dermcidin (DCD)	1	-1,42425	-3,64133	0,002	0,190	0.82
P31944	Caspase-14 (CASP14)	2	-1,22511	-2,76569	0,001	0,117	0.81
P51571	Translocon-associated protein subunit delta (SSR4)	1	-3,32742	-5,05475	0,000	0,053	0.93

Q5D862	Filaggrin-2 (FLG2)	1	0,23293	-2,88369	0,000	0,032	0.84
P17900	Ganglioside GM2 activator (GM2A)	2	-0,23326	-1,27512	0,007	0,420	0.78
P38646	Stress-70 protein (HSPA9)	2	0,78660	-2,22214	0,000	0,000	0.90
P84243	Histone H3.3 (H3F3B)	1	-3,45806	-4,77885	0,011	0,437	0.71
Q96GD3	Polycomb protein (SCMH1)	1	1,09038	-2,84126	0,000	0,000	0.90
P36957	Dihydrolipoyllysine-residue succinyltransferase (DLST)	1	2,47004	-0,94362	0,000	0,000	0.90

EPR- endoplasmic reticulum; **Log2Med** – average normalized relative abundance for early (e) and advanced (adv) OSCC samples (n=10, per group); **FDR**- false discovery rate; **AUC**- area under the curve

3.4.7.1 *Functional analysis*

The String protein-protein interaction map revealed poorly connected networks containing 12 proteins and only 5 putative interconnections (Fig.67). No biological processes and/or molecular functions were found enriched by the GO annotation tool. Possible relation was depicted between SSR4 -a calcium-binding protein taking part in the regulation of the retention of endoplasmic reticulum (ER) resident proteins, CTSD – an acid protease active in intracellular protein breakdown, FLG2 – a member of the S100 protein family, and CASP14 a non-apoptotic caspase involved in epidermal differentiation (regulates maturation of the epidermis, proteolytically processing filaggrin – FLG2).

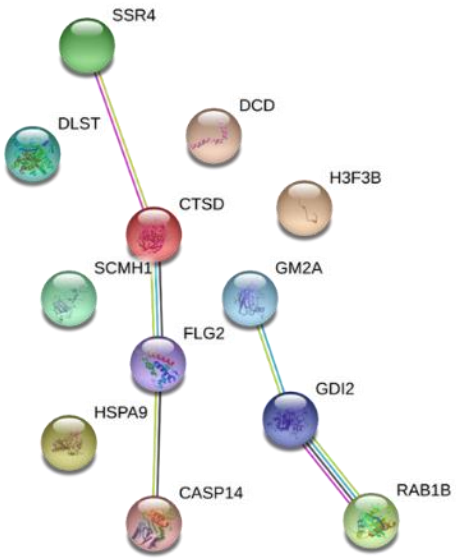


Figure 67. Functional interaction network of 12 highly expressed proteins in early OSCC showing 5 interconnections provided by String bioinformatical recourse.

V. *DISCUSSION*

Oral cavity disorders are among the most appealing ones for saliva diagnostics researchers due to their direct contact with oral tissues, and an array of detectable molecules reflecting an individual's physiological state. Sometimes, saliva may be preferred to blood and other body fluids, offering a less invasive, easily accessible method that can be performed in any setting. If becoming a routine, salivary testing for OSCC would comprise a suitable tool for population screening, assessing the treatments' outcome, monitoring of patients at high risk of recurrences, and ultimately leading to better survival (189,227). Biomarkers in saliva can reveal a carcinogenic process as well can be used to monitor its progression/remission (186). A novel focus of research is the identification of salivary biomarkers for the early detection of OSCC. Many investigators have dealt with the saliva composition of OSCC patients and have tried to find out differences when compared to normal controls. Additionally, relationships of high-risk factors such as preceding oral premalignant disorders and HPV infections with the process of oral carcinogenesis have also been widely studied (11,14,20). Sputum-derived proteins, mRNA, enzymes, and metabolites have been found at sufficiently discriminative levels between OSCC and control samples to be considered as potential biomarkers (191,228). Nevertheless, the search for specific and sensitive molecular factors with diagnostic means is ongoing, involving large-scale studies optimizing for many confounding factors introduced by the complexity of the oral environment. The current research combined state-of-the-art multidisciplinary techniques for biomarker screening of inflammatory elements, proteome, and *N*-glycome profiles in the saliva of patients with premalignant lesions (homogeneous and proliferative verrucous leukoplakia), at early and advanced OSCC stages, and healthy controls. The suggested candidates, outlined in this discovery phase, merit further investigation as potential diagnostic tools for OSCC and associated lesions.

1. *Salivary inflammatory factors as diagnostic tools for OSCC*

Since inflammation has been linked to the pathogenesis of various diseases of the oral mucosa research to date indicates the possibility of using salivary pro and anti-inflammatory factors as screening tools for those patients. It is now well recognized that altered cytokine responsiveness is tightly associated with the development of oral cancer as well as detected in patients with premalignant lesions such as oral leukoplakia (171,176,229). The commercially available multiplex-bead-based assay demonstrated to

be useful for the simultaneous evaluation of salivary concentrations of several biomarkers as part of broad profiling of immune responses. In this study, we measured salivary concentration of 8 analytes including IL-1α, Il-6, IL-8, TNF-α, IP-10, MCP-1, HCC-1, and PF4 to estimate variations in their biological activity among patients with homogeneous and proliferative verrucous leukoplakia, early and advanced OSCC stages, and normal control individuals. The analysed cases and controls were matched by age and sex. To reduce the risk of interfering variables affecting salivary concentrations of assessed cytokines, we implemented strict exclusion criteria. The analysis of the standards showed technical stability in the detection of the cytokines, with all of the % CV values falling below 20%. According to the obtained quantitative data, cytokine expressions showed low variability values (Table 7, Supplementary tables) between the cytokines as well as between the studied groups. This aspect of variability is very important for inferential statistics where relatively small samples are used to answer questions about populations. PCA outliers were spotted, especially in the advanced OSCC group. However, they were considered in the further analysis, given that PCA seems to reflect an increase in the variability of the cytokine values and at the same time of these values of greater variance. The results of our study support previous researches indicating that inflammatory elements including IL-6, IL-8, TNF-α, MCP-1, HCC-1, and PF4 may play an important role in the early detection of OSCC (169,170,230,231). Alterations in host immunity, inflammation, angiogenesis, and metabolism have been noted as the prominent pathological features in patients with oral cancer. NF-kappa-dependent cytokines are molecular messengers highly involved in all these processes (169,230,232) the altered biological activity of which have been reported not only in patients with OSCC but also with premalignant lesions, such as oral leukoplakia (OL) (170,232). Comparative analysis of patients with homogeneous (HL) (Fig.26) and proliferative verrucous leukoplakia (PVL) (Fig.27) and matched controls showed significantly elevated salivary IL-6, IL-8, TNF-α, MCP-1, and HCC-1 in the current investigation. Increased salivary levels of the first three mediators have previously been reported in HL and other pre-neoplastic lesions by various authors (170,172,173,233,234), while a single study has demonstrated higher IL-6 in PVL saliva collated to controls (235). In our experiments, IL-1α levels did not differ among the two clinical forms of leukoplakia and controls, in contrast to Rhodus *et al.* (170) reported its elevated expression of this cytokine in dysplastic oral premalignant lesions. To the best of our knowledge, we are the

first to estimate salivary MCP-1, IP-10, HCC-1, and PF-4 in HL, together with IL-1α, IL-8, and TNF-α in PVL. Although IP-10 and PF-4 levels were also observed increased in both leukoplakia forms, statistical significance was not confirmed. However, immunohistochemical studies of paraffin-embedded tissue sections from oral lesions presenting epithelial dysplasia showed intense PF-4 (CXCL14) staining (236). Our study also found a positive correlation of altered HCC-1 and IL-6 concentrations with lesions' histologic features (Fig.35), suggesting them as putative markers for epithelial dysplasia. Similarly, Sharma *et al*. (237) described that within the leukoplakia group IL-6 level was seen to increase with the severity of dysplasia grade. Histological grading of OL has been proposed to influence salivary cytokine expression in a growth manner, correlating with the level of epithelial dysplasia (172). Collectively, these results indicate that proinflammatory cytokines, especially IL-6, IL-8, TNF-α, MCP-1, and HCC-1 could be importantly involved in the pathogenesis of OL and its manifestation forms. Inflammatory responses, as well as alterations in immune system of OSCC patients, play a pivotal role in disease progression (175). The relation between OSCC and chronic inflammation has been proved by the imbalance in local and systemic immunomodulatory cytokine levels (238,239), leading to tumor growth and proliferation. In the present research 8 salivary cytokines were estimated in patients with OSCC, at early (I&II) and advanced (II&IV) clinical stages. Six biomarkers with significantly higher expression in early OSCC than in controls were found: IL-6, IL-8, TNF-α, MCP-1, HCC-1, and PF-4 (Fig.29). Among them, IL-6 and TNF-α marked a considerable growth towards the OSCC evolution (Fig.31) indicating the potential involvement of these mediators in disease progression and severity. Cytokines are implicated in different disease states and are dependent upon a variety of pathways. Their functions are multifaceted including inflammation (e.g. IL-1α, IL-6), and inflammation combined with an immune response (e.g., IL-10, TNF-α) (168). IL-6 has been shown to activate STAT3 oncogene which in its turn activates VEGF and tumor angiogenesis (240). Rising IL-6 levels were observed from well to moderate and to poorly differentiated OSCC tumors in serum and saliva suggesting that this cytokine can be associated with the severity and aggressiveness of the disease (234). IL-8 (CXCL8) is one of the dominant transcriptional targets of the inflammatory signaling mediated by NF-κB, which is commonly activated in cancer cells. It is a proinflammatory chemokine that acts on leukocytes and endothelial cells, via their receptors, to promote immune infiltration and angiogenesis, which in turn

establishes a venue for cancer cell local invasion, migration, and metastasis (241). TNF-α, a widely expressed cytokine during oral cancer transformation and progression has been reported to be involved in various OSCC pathogenic pathways (242). Our results appeared to be consistent with previous findings in which increased levels of NF-κB associated IL-6, IL-8, and TNF-α in oral cancer saliva have been reported, suggesting that OSCC progression is likely enhanced by continued expression of proinflammatory, pro-angiogenic cytokines (175,230,243,244). Nevertheless, there are a few researches describing salivary cytokine levels at different OSCC clinical stages. Lee *et al.* (245) found significantly upregulated salivary IL-6, IL-8 and TNF-α levels in early OSCC stages (I/II) than in control subjects, but no distinctive activity between early and advanced (III/IV) stages. Similarly, Dineshkumar *et al.* (234) stated no significant difference in salivary IL-6 based on OSCC clinical staging. Krishnan et al. (242) revealed important TNF-α overexpression in sputum of OSCC patients at stage IV compared to all other stages. OSCC driven modulation of IL-1α remains controversial. According to our results, no differential IL-1α levels were seen between controls, early and advanced OSCC stages, in agreement with the observations of Lee et al. (245) and Babiuch *et al.* (232). In contrast, other authors described elevated IL-1α in OSCC saliva (170,175,244). The transcription of angiogenic and tumourigenic chemokine genes is modulated, in part, by the NF-κB family of transcription factors (246). Chemokines are secreted in response to signals such as proinflammatory members of the IL-1 family, tumor necrosis factors, and interferon-c (IFN-c) and thus playing an important role in selectively recruiting monocytes, neutrophils, and lymphocytes (231). The complex interaction among these immune modulators in OSCC have been presented by Pearson's correlation analysis displaying significant positive interconnections among the explored salivary cytokines (Table 8D and E, Appendix). Overall, altered chemokine function in cancer promotes cell survival, enhances proliferation, neovascularization, motility and metastasis in multiple tumour types (247). Several studies have implicated a number of chemokines and their receptors in squamous cell cancers of the head and neck arguing that tumor-related changes in chemokine composition are detectable in oral fluid (248,249). Indeed, our results showed considerable overexpression of inflammatory MCP-1 along with homeostatic HCC-1 and PF-4 in the saliva of patients at early OSCC stages compared to normal controls. Monocyte chemoattractant protein 1 (MCP-1) or CCL2, regulates monocyte migration by promoting their circulation to the site of inflammation (250).

Besides, it is frequently expressed as tumor cell-associated chemokine. Increased MCP-1 in OSCC and metastatic lymph nodes have been detected by Ferreira and co-workers (251) while its expression in HNSCC was associated with tumor invasion in oesophageal SCC (252). HCC-1 or hemofiltrate C-C motif chemokine 14 (CCL14) is a homeostatic chemokine found to promote angiogenesis and tumor progression (253). Its involvement in oral carcinogenesis has been reported by Feng *et al.* (254) exhibiting differentially expressed long non-coding RNA (lncRNA) HCC-1 transcripts detected in oral mucosal samples from OSCC patients. Platelet factor 4 (PF-4) or CXCL14, also known as BRAK is a highly conserved homeostatic chemokine. Responsible for immune cell recruitment and maturation, as well as influencing epithelial cell motility, PF-4 contributes to the establishment of immune surveillance within normal epithelial layers. Previously, ELISA estimated salivary IL-8,IP-10 and HCC-1 were found notably elevated while MCP-1 marked insignificant growth in patients with oral and laryngeal squamous carcinoma before therapy compared to subjects with periodontitis (249). The authors presumed that the impact of periodontal infections on the salivary chemokine composition is probably limited due to similar IL-8 concentrations of pre-and post-treatment periodontitis and healthy subjects. However, ELISA quantification resulted in very low detection levels and more refined methods could likely indicate not only intact chemokines, but also those modified post-translationally. Despite the complex functional characteristics, PF-4 is thought to be a key regulatory factor in cancer. Dysregulated PF-4 was found in several carcinoma types including head and neck neoplasia. Its disruption was shown to limit critical antitumour immune regulation and to correlate to poor patient prognosis (255). Importantly increased salivary PF-4 concentration in OSCC was firstly described in the current study. In contrast, PF-4 expression has been shown remarkably decreased in OSCC cells and its induced up-regulation resulted in suppressed activity toward tumour progression of oral cancer in vivo (256,257). Frederick et al. (258) disclosed *In situ* hybridization analysis where BRAK (PF-4) mRNA expression was detected in normal and dysplastic tongue but absent in buccal and metastatic OSCC of the tongue. Analogous, Nakayama and colleagues (259) described significantly decreased PF-4 mRNA levels in primary and metastatic OSCC compared with tumour free tissue based on QRT-PCR analysis. Furthermore, both groups found numerous inflammatory and stromal cells, showing PF-4 expression predominantly in the cytoplasm in adjacent non-tumour tissues but significantly down regulated in OSCC tissues. The

present outcomes revealed higher IP-10 concentration within the OSCC than in the control group, yet, not statistically validated, which is in a line with a recent publication wherein no alterations of IP-10 in saliva of patients at early OSCC stages have been estimated (245). IP-10 (CXCL10) is an angiostatic chemokine that mediates biological activity through binding G protein coupled receptors and glycosoaminoglycans (249). Although the versatile properties of IP-10, it appears to be secreted mainly by tumour cells and therefore might be a powerful biomarker in cancer progression. Significantly elevated IP-10 levels were detected in serum samples from HNSCC patients (260) as well as in tissue and peripheral blood of individuals with nasopharyngeal carcinoma (261). Besides, IP-10 has been suggested as a potential marker for response to radiotherapy and overall survival in patients with tongue SCC (262). A common way of evaluating diagnostic accuracy of individual biomarkers is by using Receiver Operating Characteristics (ROC) analysis, which plots the true positive rate versus the false positive rate of a particular disease (263). The discriminatory efficacy of the studied markers to discern early OSCC form individuals without oral lesions (controls) was displayed with ROC curves (Fig.29). IL-6, IL-8, TNF-α and HCC-1 had AUC values higher than 0.8, the former three being consistent with the results of Lee *at al.* (245). Superior sensitivity of IL-6 and IL-8 in detecting OSCC has also been evaluated in two large scale studies (230) and (234), respectively. MCP-1 and PF-4 had AUC higher than 0.7, while IL-1α and IP-10 did not show sufficient sensitivity of specificity for early OSCC diagnosis (AUC = 0.5 and 0.6, respectively). This discovery suggests that a combination of biomarkers might increase their clinical utility. According to the obtained correlation analysis, no association have been identified between altered cytokine responsiveness and OSCC patient's clinicopathological variables such as sex, smoking habits, tumour oral location, clinical type and size. It does not mean that a potential relationship does not exist, but that the sample cohorts were not large enough to confirm statistical significance. However, altered TNF-α expression has significantly been correlated to OSCC (Table 8) and (Fig.37), being the marker with the highest ROC AUC values of 0.94 and 0.95 in early and advanced disease stages compared to controls (Fig.29D) and (Fig.30D, Supplementary figures), respectively. TNF-α and its soluble receptors are strongly suggested to be beneficial in detection, staging and/or predicting prognosis of oral cancers (242). A significant increase of salivary TNF-α in advanced OSCC denotes its involvement in accelerating the disease progressions. Our observations are corroborating the ones from

Jablonska *et al.* (264) wherein a relation between clinical staging and increased serum TNF-α levels in OSCC has been discovered. The positive correlation between IL-6 and the presence of cervical adenopathy in oral cancer patients found in this study substantiate earlier disclosures. Riedel *et al.* (265) and Tartour *et al.* (266) revealed statistically significant association between positive lymph nodes in HNSCC and elevated serum IL-6 level. It is now well established that OSCC may occur through evolution of oral potentially malignant lesions (267). There are numerous investigations that have assessed and compared cytokine expressions in body fluids, tissue specimens and cell culture models of OSCC and OPMDs. However, a few of them have considered more than one cytokine when comparing different premalignant lesions. One of the objectives in this study was to estimate differences between the target cytokine expressions in HL, PVL and OSCC patients. Statistical tests resulted in distinctively increased salivary IL-6, IL-8, TNF-α, HCC-1, and PF-4 in OSCC, independent on the clinical stage, compared to individuals with HL and PVL (Fig.32 and 33, Supplementary figures). All of the analytes listed above, except PF-4, were notably higher in HL and PVL than in subjects without oral lesions (controls). Whereas the premalignant microenvironment elicits proinflammatory cytokine production, the tumour microenvironment seems significantly more immune stimulatory. It could be assumed that malignant transformation of HL and PVL may be influenced by the continuous exposure to the overexpression of these proinflammatory, proangiogenic cytokines. The fact that the same molecules were remarkably elevated both in OSCC and OL may have a diagnostic significance for cancerous conversion of potentially malignant lesions. A few studies have addressed the diagnostic potential of routine cytokine measurements in screening of OSCC. Rhodus *et al.* (170) reported significantly higher levels of IL-1α, IL-6, IL-8, and TNF-α in OSCC saliva as compared to oral premalignant lesions with moderate and severe dysplasia. In accordance are the revelations of other authors, stating markedly elevated IL-6 (234), IL-8 (268,269) and TNF- α (242) in saliva of subjects with OSCC in comparison with OL and normal controls. Pronounced increase of salivary HCC-1 and PF-4 in early OSCC stages, compared to HL and PVL was primary characterized in this study, bringing new insights into the search of potential biomarkers for early OSCC diagnosis. Further, direct association has been found between IL-6, IL-8, TNF-α, MCP-1, HCC-1, PF-4 and IP-10 levels, OSCC and OL (Fig.37). However, cytokine activity showed to be influenced between each other, highlighting their intricate interactions. IL-6 being not

affected by any of the investigated modulators might have an independent role in the regulation of the inflammatory response in OSCC and associated lesions. It has been demonstrated that inflammatory factors including cytokines and chemokines are abundantly expressed in the course of periodontitis (270,271). With regards to cancer risk, chemotactic cytokines and chemokines are of particular interest as they are involved in periodontal pathogenesis but also expressed in healthy sites (272). To avoid any potential interference, our cohorts of volunteer donors had been pre-screened to eliminate those with acute or chronic inflammation in the oral cavity upon direct clinical visual exploration. Nevertheless, considering that local inflammatory conditions may result in cytokine overexpression, investigators have found the imput of OSCC to the elevation of these molecules to outweigh any potential background conferred by the host's inflammatory condition (271).

2. *Salivary N-glycans as biomarkers for OSCC*

Approaches for glycan analysis have evolved with the development of high-throughput platform technologies allowing the evaluation of large sample cohorts in an efficient manner. Among these methodologies, LC combined with MS is a flexible and powerful tool for glycomics study. It enables the detection of structures with increased dynamic range, pushing lower the detection limit of analysis in complex samples like plasma, serum, urine, and even cerebrospinal fluid (24,191). Besides, MS can determine the molecular mass of analytes and clarify *N*-glycans structure and composition. Novel research revealed the biological significance of the serum *N*-glycome in human health and disease, focusing on the exploration of the connection between altered *N*-linked glycosylation and various pathologies, placing a special emphasis on cancer. The possibility to differ protein glycosylation patterns between cancer and control patients highlights glycobiology as a potential area for biomarker discovery. The use of serum glycans as detecting and monitoring agents have been reported in different neoplasia including colon, prostate, pancreatic and ovarian cancers (273). Recent studies have demonstrated the potential of salivary glycopatterns to serve as diagnostic tools for gastric and breast cancers (274,275). Oral cancer-associated aberrant glycosylation has been proposed by Lin et al. who found that acetylgalactosaminyltransferase 2 (GALNT2) was overexpressed in 73% of the examined OSCC tissue specimens (209). Moreover, OSCC driven molecular mechanisms

were suggested to be due to dysregulated cellular networks among DPAGT1 (*N*-glycosylation-regulating gene), E-cadherin adhesion, and canonical Wnt signaling in a cultured cell line (276). The current study describes a non-invasive approach for testing saliva as an alternative to blood serum source of putative glycan biomarkers for early OSCC diagnosis. Furthermore, this is the first report on salivary *N*-glycan profiles identified in patients with the oral potentially malignant disorder (PVL) and different OSCC stages compared to healthy controls to determine oral cancer-associated *N*-glycosylation changes. The substantial division of the *N*-glycan profiles of individuals with and without oral lesions seen on the PCA biplot suggests variability of salivary glycosylation patterns, possibly related to disease-dependent *N*-glycomic aberrations. Besides, we suggest that differences are rather quantitative than qualitative which is in accordance with other studies. For initial glycan identification, UPLC was coupled with MS for confident assignment of monosaccharide composition. The chromatographic separation facilitates the reduction of sample complexity and enhances dynamic range. HILIC empowered high-resolution separation of isobaric structures, together with MS analysis, resulted in multiple components, confidentially identified with mass accuracies lower than 20 ppm. The intricate and highly abundant glycan population in saliva has been formerly described reporting different numbers of identified oligosaccharides. Depending on the sample size and saliva type, published studies reported 60 *N*-glycans in parotid saliva (277), 78 (278), and 265 (279) in the whole saliva. In this study, we listed 90 compounds detected in the whole saliva of individuals without oral lesions (controls). The majority of them are neutral compounds, with high-mannose and complex bi-, tri-, and tetra-antennary structures, emphasizing the heterogeneous and complex glycan population in sputum. We also observed that some compositions were present in multiple peaks implying structural isomerism. From the putative monosaccharide, compositions of the described *N*-glycans could be seen that the majority are core/antenna fucosylated, bi, and tri- antennary structures. Fucosylated epitopes such as Lewis (Lex, Ley, and SLex) and H-type antigens were present as well, confirming the findings of Everest-Dass *et al*. (278) and Guile *et al* (277). Overexpression of these fucose-containing moieties frequently occurs on the cancer cell surface and has been associated with tumor progression and metastasis (280). Ley carbohydrate up-regulation has been demonstrated in OSCC cell lines and correlated to the poor prognosis of oral cancer patients (281). Aberrant fucosylation is related to

disruption of cell-cell interactions and changes of growth factor receptor functions, eventually redounding to metastasis (282). Elevated fucose and fucosylated serum protein levels of patients with OSCC tumors have previously been demonstrated (283,284). In the present study, six out of the eight glycans showed considerably different expression in OSCC were fucosylated. Three of them observed at *m/z*=1045.89, *m/z*=1134.91, and *m/z*=1389.50 (Fig.40A, B, and C, respectively) marked significantly decreased relative abundance compared to PVL patients and controls. The decreased relative intensity of fucosylated bi-antennary *N*-glycans can also be observed in oesophageal adenocarcinoma serum samples (285). The reduced expression might be related to the carcinogenic process, being notable at early and advanced OSCC stages. The other three fucosylated *N*-glycans noticed at m/z = 842.81, m/z = 915.3, and m/z = 717.27 (Fig 41A, B, and C, respectively) exhibited significantly elevated relative abundance in OSCC versus patients with premalignant lesions (PVL). Increased levels of these glycans at different OSCC stages suggest for *N*-glycosylation changes, which has occurred due to the evolution of an existing oral dysplastic lesion into tumour formation. Highly expressed fucosylated *N*-glycans detected in saliva from OSCC patients in the current research is in accordance to previous reports where elevated fucose and fucosylated proteins were detected in OSCC serum and cell line models (283,286). Furthermore, Shah *et al* suggested *L*-fucose as a candidate OSCC serum biomarker due to its notable, gradual increase from control, to premalignant disorders to oral cancer samples (284). Mono and di-sialyated structures were characterised in the total *N*-glycan pool investigated. In carcinogenesis, altered expression of sialic acid (SA) is suggested to be involved in malignant transformation and progression (287). Study from 2017 reported significantly higher expression of SA on outer cell membranes, serum and saliva obtained from patients with premalignant lesions and OSCC when compared to healthy subjects (288). Mono-sialylated core fucosylated glycan observed at m/z= 915.33 (Fig. 41C) was notably higher in OSCC compared to the PVL group, evoking speculation for cancer-related glycosylation change, preceded from OPMD. The distinctive expression of salivary *N*-glycans between patients with oral cancer and associated lesions highlights their potential as non-invasive biomarkers that can provide insights into disease state and progression. However, the isomeric and branched nature of oligosaccharides and the complicated linkage are the rationale for the defiant assessment of exact structures. To overcome this challenge, multidimensional, high throughput

analytical approaches have proved to be efficient for robust, reproducible, and large-scale sample analysis, fundamentally required in the biomarkers discovery pathway.

3. *Salivary proteome profiling of OSCC and associated lesions*

The cellular and molecular heterogeneity of OSCC and the large number of molecular changes involved in oral carcinogenesis emphasize the importance of studying proteins responsible for them by global scale proteomics. To compile a list of candidate biomarkers for OSCC detection we carried out a comprehensive two-step approach using saliva as a diagnostic biological fluid. In the first non-targeted label-free LC-MS step, a spectral library was built, comprised of all the proteins identified in the OSCC salivary pool. Secondly, quantitative profiling of proteins expressed in healthy control, PVL, early and advanced OSCC individuals was generated by data-independent LC-MS/MS (SWATH) analysis (289). To generate a list of putative biomarker panel(s), we used logistic regression (LR), discriminant analysis, analysis to process the results obtained from the 691 quantified proteins in the four sample cohorts. These statistics represent a measure of the utility of the appointed proteins to differentiate the healthy and diseased states. Penalized multinomial LR selected 49 proteins with altered expressions, (Table 17, Supplementary tables) partially classifying the samples according to their group of belonging. The majority of these differential proteins were further elected by binomial LR (two-by-two comparisons) which demonstrated the ability to clearly order the samples so thus to discriminate between the investigated groups. Statistical significance of changed protein abundance is shown by Student's t-test FDR adjusted *p*-values from pair-wised analysis of controls versus patients at early and advanced OSCC stages (Table 19 and 20, respectively). The comprehensive catalog of the human salivary proteome exceeds more than a thousand annotated proteins, reflecting the diversity of molecules available for biomarker screening (290). To obtain more insights into the modulations related to OSCC, an extensive literature search was carried out to see which proteins have been associated with oncogenesis. Most of the markers exhibited differential expressions in the current study were already related to cancer, including OSCC, and 23 proteins were identified to be present in saliva in this pathological condition. Proteins exhibited significantly altered

abundance in OSCC in our results, coinciding with the findings of other authors are shown in Table 25.

Table 25. List of proteins differentially expressed in OSCC compared to controls, detected in the current and other oral cancer-associated studies.

Protein name	Expression	Sample	Comment	Reference
C3, SERPINA1, TYMP	up-regulated	saliva	correlate with the risk of developing OSCC	(194);(291)
C3, SERPINA1, HPX	up-regulated	saliva	suggested as potential OSCC biomakrers	(188);(292); (293)
TTR, ORM1, ORM2, CST4	down-regulated	saliva	suggested as potential OSCC biomarkers	(188)
APOA1, ITIH1, protein AMBP, VTN, RPL7	up-regulated	saliva	overexpressed in OSCC	(294);(295)
PLG, AGT (SERPINA8)	up-regulated	saliva, tissue	found to be a predictive marker for early OSCC; associated with malignant epithelial neoplasia including OSCC	(295);(296); (297)
SEPRING1, HP, ORM1	up-regulated	saliva, plasma	identified as saliva and plasma markers of OSCC	(295);(298)
RAB1B, SERPINB4	up-regulated	saliva, tissue	SerpinaB4 identified in OSCC saliva; both found overexpressed in gingival SCC tumors	(299)
S100-A7	up-regulated	saliva, tissue	increased in T1 and T2 OSCC; successfully discriminating HNSCC from normal tissues; up-regulated in early, non-metastatic versus advanced, metastatic OSCC	(300);(301); (302)
CALML5	disregulated	tissue	CALM5 suggested as a candidate cancer driver gene in OSCC	(303)
CASP14, FLG	down-regulated	tissue	partial loss in invasive oral carcinomas	(304)
PADI4	up-regulated	cells	suggested to stimulate esophageal SCC tumor growth *in vivo*	(305)
WARS	up-regulated	tissue	overexpressed in OSCC tissues suggested to enhances oral cancer cell invasiveness	(306)
SERPINA3	up-regulated	tissue	found upregulation in lymph node positive OSCC patients	(307)
HMGB2	up-regulated	cells	overexpressed in a panel of HNSCC cell lines based on mass spectrometry analysis	(308)
LYZ	up-regulated	saliva	identified in saliva OSCC samples by shotgun proteomics	(295)
CDH1,VIM, PFN1	down-regulated	tissue, cells	CDH dysfunction promotes tumour progression;correlation between decreased CDH1 and VIM expression and aggressive HNSC; decreased PFN1 in advanced OSCC	(309);(310); (311) (312)(313)
CAT	up-regulated	saliva	Increased CAT levels in OSCC patients compared to healthy controls	(314)
FKBP1A (Pin1)	up-regulated	cells	Pin1 overexpression is associated with lymph node metastasis through cyclinD regulation in OSCC.	(315)
SQRDL	indifferent/upregulated	tissue	examined by western blotting in OSCC and compared to benign tissue	(316)
TUFM	down-regulated	saliva	detected by quantitative proteomics	(294)

ZG16B / PAUF	down-regulated	tissue, cells	PAUF expression was correlated with nodal metastasis and poor prognosis in OSCC	(317)
MUC7	down-regulated	tissue	a gradual decrease in acidic mucin intensity in OED and overall weak intensity in OSCC.	(318)
ARPC5	up-regulated	tissue, cells	Highly expressed in HNSCC tissues; Silencing of ARPC5 inhibited cell migration and invasion in HNSCC cell lines.	(319)
HSPA9 (HPS70)	up-regulated	tissue	overexpressed during tumourigenesis of the squamous epithelium of the oral cavity	(320)
CTSD	up-regulated	tissue	CTSD expression correlated with OSCC invasion and progression	(321)
H3.3	differential	saliva	H3F3A transcripts elevated in OSCC patients; downregulated in OSCC compared to controls	(322);(294)

Complement C3 (C3); Alpha-1-antitrypsin (SERPINA1); Thymidine phosphorylase (TYMP); Cystatin-S (CST4); Hemopexin (HPX); Transthyretin (TTR); Alpha-acid glycoprotein 1 (ORM1); Alpha-1-acid glycoprotein 2 (ORM2); Apolipoprotein A-I (APOA1); Inter-alpha-trypsin inhibitor heavy chain H1 (ITIH1); Protein (AMBP);Vitronectin (VTN); 60S ribosomal protein L7 (RPL7); Plasminogen (PLG); Angiotensinogen (AGT/Serpin A8); Plasma protease C1 inhibitor (SERPING1); Haptoglobin (HP); Ras-related protein (RAB1B); Leupin (SERPINB4); Protein S100-A7 (S100A7); Calmodulin-like protein 5 (CALML5); Caspase-14 (CASP14); Filaggrin-2 (FLG2); Protein-arginine deiminase type-4 (PADI4); Tryptophanyl-tRNA synthetase (WARS); Alpha-1-antichymotrypsin (GIG25/ SERPINA3); High mobility group protein B2 (HMGB2); Lysozyme C (LYZ); Cadherin-1 (CDH1); Vimentin (VIM); Vinculin (VCN); Profilin-1 (PFN1); Catalase (CAT); Peptidyl-prolyl cis-trans isomerase (FKBP1A); Sulfide:quinone oxidoreductase (SQRDL); Elongation factor Tu (TUFM); Zymogen granule protein 16 homolog B (ZG16B); Mucin-7 (MUC7); Actin-related protein 2/3 complex subunit 5 (ARPC5); Stress-70 protein (HSPA9); Cathepsin D (CTSD); Dermcidin (DCD); Histone H3.3 (H3F3B).

Our revelations are compatible with other proteomic studies where differentially expressed proteins have been proposed as candidate OSCC biomarkers. For instance, C3, SERPINA1, TYMP, HEM, APOA1, ITIH1, protein AMBP, VTN, SERPING1, SERPINB4, HP, ORM1, and S100-A7 were found at significantly increased abundance in OSCC compared to control saliva (188,194,294,295,298,300,323). In addition, RAB1B, WARS, SERPINA3, and SQRDL were shown up-regulated in OSCC tissues and/or cell culture and associated with positive lymph nodes, and cancer invasiveness (299,306,307,316). We showed that caspase 14 (CASP14) and filaggrin (FLG) exhibited higher abundance at early compared to advanced cancer stages. Scharenberg et.al (95) reported co-localization of the two proteins in formalin-fixed OSCC tissues where CASP14/FLG expression is transmitted from the primary tumor to its metastasis and partial loss of both proteins was seen in invasive oral

tumors. Similar to our findings, angiotensinogen (SERPINA8) and plasminogen (PLG) upregulated expression in OSCC saliva was found by Csősz and co-workers (295). Bioinformatics analysis has shown that the plasminogen activator system could serve as a predictive marker for early OSCC (86), while angiotensin-converting enzymes were associated with malignant epithelial neoplasia such as OSCC (87). *CALML5*, encoding a skin-specific calcium-binding protein was found recurrently mutated in OSCC and suggested as a novel candidate cancer driver gene in oral carcinogenesis (303). High CTSD expression was closely correlated with oral carcinoma invasion and progression (112). However, some proteins were shown to have contradictive our findings expressions. Independent studies demonstrated RPL7 (294) and LYZ (295) upregulation, while TUFM and ORM2 downregulation was estimated in OSCC collated to saliva obtained from healthy volunteers (188). Besides, CST4, PADI4, HMGB2, FKBP1A (Pin1), ARPC5, HSPA9 exhibited decreased expression in OSCC in our analysis, while other authors found them overexpressed in OSCC tissue and/or cultured HNSCC cell lines (305,308,315,319,320). Such a discrepancy may be explained with inter-person and/or sample type variabilities. Meanwhile, some underexpressed salivary proteins have shown similar expression profiles in other OSCC specimens and/or cancer types. For instance, CDH1, VCL, VIM, and PFN1 were found at a significantly lowered relative abundance in OSCC compared to non-cancerous tissue samples. Loss of cell-cell adhesion is important for the development of cancer invasion and metastasis (324). CDH1 or epithelial cadherin (E-cadherin) is involved in mechanisms regulating cell-cell adhesions, mobility, and proliferation of epithelial cells. It has long been recognized as a suppressor of invasion and metastasis in many contexts, so thus its depletions is suggested to be an important determinant of tumour progression (309). Furthermore, a hallmark of epithelial-mesenchymal transition (EMT) is the loss of E-cadherin which may play an important role in OSCC carcinogenesis or progression. Downregulation of E-cadherin expression has previously been demonstrated in HNSC cell lines (310) and oral cancer tumor tissue (311). Vinculin (VCL) is a key adhesion-related protein that affects tumorigenesis, metastasis, and invasion. It has been shown that cells with low VCL expression have increased metastatic potential and are closely related to the degree of tumor progression in non-small cell lung cancer (313). Other studies have observed similar phenomena in the breast (325) and colon (326) cancers, suggesting VCL

as a prognostic marker and a potential target for treatment. Vimentin (VIM) is a type III intermediate filament protein involved in cell attachment, migration, and signaling. The expression of mesenchymal markers such as VIM is observed after EMT, an important initial behavioral change related to the adhesion and migration properties of abnormal cells that are required for local tumor invasion. Immunophenotypic analysis of primary tumor tissues demonstrated a significant correlation between decreased E-cadherin and VIM expression and aggressive cancer features in HNSCC patients (310). VIM overexpression examined by immunohistochemistry was closely associated with local recurrence and short disease-free survival in OSCC, while tumours lacking or expressing a low level of VIM were correlated with a better prognosis (311). Profilin1 (PFN1), an indispensable and ubiquitously expressed actin-binding protein, has roles in normal cell motility, proliferation and differentiation. In support of our results, Adami *et al*. (312) showed lowered PFN1 expression in late stage OSCC tumour tissues. In addition, Peng and co-workers have reported association of PFN1 reduction with advanced laryngeal SCC tumours and a trend for enhanced lymph node invasion. No previous data were found on GLUL, GM2A, TRIM29, IGHG2, Polycomb protein SCMH1 and SPARCL1 association with OSCC, however, their modulated expression have been related to other cancer types (132–136). Apart from potential clinical applications, these differential proteins may contribute to better understand the molecular mechanism of the disease. While the salivary proteome of the head-and-neck cancer has been widely explored, a few studies have described salivary proteome profiles of patients with OPMD and compared them to those with oral cancer (294,327). The investigated OPMD cohorts comprised of patients presenting a variety of oral lesions such as erythroleukoplakia, oral submucous fibrosis (OSF), homogeneous, speckle and verrucous leukoplakia, verrucous hyperplasia, lichenoid lesions, etc. whereas our research have particularly focused on PVL due to its high risk of malignant transformation (29–31). To reveal the representative feature of OSCC salivary proteome, we attempted to distinguish OSCC from PVL by using the abundance of the salivary proteins. Our LR selected a series of differential proteins that could discriminate between the two pathologic states, the statistical significance of which was further pair-wisely assessed (Table 26). A literature review disclosed that many of them have previously been described in studies regarding OSCC and associated lesions.

Table 26. List of proteins differentially expressed in OSCC compared to PVL detected in the current, and other oral cancer-associated studies.

Protein name	Sample	Expression	Comment	Reference
APOA1, HP, HSPA5	saliva	up-regulated	offers a clinically effective tool for detecting OSCC and monitoring OPMDs	(294);(327)
RAB1B, FGA, C3, HPX, SERPING1, TYMP, PLG	saliva, blood	up-regulated	Biomarker candidates of OSCC	(295);(298); (299)
KLK1	tissue	up-regulated	abundantly expressed in human OSCC and may be implicated in malignant progression	(328)
SOD2	tissue	up-regulated	in premalignant and OSCC lesions early event in oral carcinogenesis	(329)
LRG1	saliva	up-regulated	candidate biomarkers of OSCC	(194)
KNG1	saliva	up -reglated	higher in OSCC than in OPMD; candidate biomarkers of OSCC	(294);(306); (327)
HIST2H2AA	saliva	up-regulated	higher in OSCC than in OPMD	(294)
DSP,PKP1, DSC2	tissue	down-regulated	decrease immunoreactivity in dysplastic and OSCC tissue samples	(330);(307)
RHCG	tissue, cell lines	down-regulated	RHCG acts as a tumor suppressor gene that plays a crucial role in inhibiting tumorigenicity and metastasis in HNSCC	(331)
KRT type II	saliva, tissue	differential	early feature in the pathogenesis of invasive OSCC; downregulated expression in OED and OSCC	(332);(291); (333)
CNFN	saliva, tissue	down-regulated	found in healthy controls; significantly downregulation in HNSCC tumor tissue	(188)

Apolipoprotein A (APOA); Haptoglobin (HP); SPARC protein; Endoplasmic reticulum chaperone BiP (HSPA5); Ras-related protein (RAB1B); Fibrinogen alpha chain (FGA); Complement C3 (C3); Hemopexin (HPX); Plasma protease C1 inhibitor (SERPING1); Thymidine phosphorylase (TYMP); Plasmiogen (PLG); Kallikrein-1 (KLK1);Superoxide dismutase (SOD2); Leucine-rich alpha-2-glycoprotein (LRG1); Kininogen-1 (KNG1); Histone H2A type 2-A (HIST2H2AA); Desmoplakin (DSP); Plakophilin-1 (PKP1); Desmocollin-2 (DSC2); Ammonium transporter Rh type C (RHCG); Keratin, type II; Cornifelin (CNFN); oral epithelial dysplasia (OED); oral potentially malignant disorder (OPMD)

HSPA5, KNG1, and HIST2H2AA are part of a protein panel that appeared to be suitable for detecting OSCC and monitoring high-risk OPMD cases in Taiwan's Oral Cancer Screening Program (327). RAB1B, C3, HPX, APOA1, HP, LRG1, SERPING1, TYMP, and PLG have been proposed as potential OSCC biomarkers, identified in human saliva (194,295,299,327). SOD2 up-regulation in premalignant and OSCC oral tissues has been suggested as an early event in oral carcinogenesis (329), and according to Pettus *et al.* (328), the abundant expression of various kallikreins in human OSCC may be implicated in malignant

progression. Some downregulated proteins found in saliva such as DSP, PKP1, DSC2, RHCG, KRT type II, and CNFN exhibited similar manifestations in histologic studies. For instance, decreased immunoreactivity of the desmosomal components DSP, PKP1, and DSC2 has been detected in dysplastic and OSCC tissue samples compared to the normal epithelium (330). Desmosomes are the most prominent cell-cell junctional complex in stratified squamous epithelial tissues. According to the authors, the loss of desmosomal adhesion is implied to be an early occurrence in the progression to SCC, connected with increased migratory capacity of tumor cells. Examination of these cell adhesion molecules may provide a good marker of increased risk for progression to OSCC. Keratins have important molecular functions as structural constituents of the cytoskeleton as well as implications on cell shape and cell size. Keratin type II overexpression in OSCC saliva has been reported by Krapfenbauer et al. (291). The regular keratin patterns in normal epithelia and their aberrant expression in premalignant lesions and carcinomas present possibilities for improving the early diagnosis of oral epithelial neoplasms (334). In this work we demonstrated the use of human saliva collected from healthy volunteers, OPMD, and OSCC patients, to measure candidate biomarkers using SWATH assay in a discovery-based proteomic study. Among the identified proteins, the statistical models indicated a panel of putative biomarkers that were differentially abundant between the studied cohorts. To highlight the molecular mechanisms potentially involved in the alteration of OSCC microenvironments, the differentially expressed proteins were applied for group-wise exploration with the String functional annotation tool. Furthermore, the enrichment analysis showed that the highest rank of biological processes involving the differentially expressed proteins in the OSCC groups was related to active regulatory mechanisms implicating immune and inflammatory responses (APOA1, HP, TTR, ORM1, ORM2, CAT, etc.) (Fig. 29 and Fig.32). The highlighted processes are consistent with the previous notion that the immune, defense, and inflammatory responses are dysregulated in the OSCC (165,335). In addition, the highest association with the molecular function was inhibition of enzymatic activity and receptor-associated protein activity (signaling) where the Serpins family of protease inhibitors, CTSD, protein AMBP and CST4 are strongly involved. Many proteases play important roles in the invasion and metastasis of cancers because of their ability to degrade the extracellular matrix barrier surrounding tumors. The salivary

protease spectrum was found to be associated with oral diseases (336). For example, the saliva of patients with OSCC contained increased numbers of proteases than those of other oral diseases and health (336).Thus, analysis of the salivary protease spectrum may be useful as biomarkers in the screening, early diagnosis, and monitoring of cancer occurrence and evolution. Besides, our study revealed deregulated activity of proteins functioning as structural constituents of the cytoskeleton and cell-adhesion molecules in oral paraneoplastic and malignant lesions such as Keratins from type II, CNFN, CDH1, DSP, DSC2, PKP1 (Fig.56 and Fig.64). Aberrant expression of cytoskeletal components is associated with impaired epithelial differentiation and organization during tumourigenesis indicating that early atypical changes of OSCC may begin on the squamous cell skeleton (129). In addition, it has been established that the initial step in the metastatic cascade is the detachment of tumour cells from the primary tumour via dysregulation of normal cell–cell and cell–matrix interactions where cell adhesion molecules mediate these interactions (330). The applied proteomics strategy allowed the identification of numerous proteins at differential levels, demonstrating to be feasible approach of revealing putative OSCC associated biomarkers. Further validation of the selected markers is required prior to their translation into clinical application. They need to be extensively certified considering that sampling efficiency for LC-MS/MS might vary from one experiment to another and some of the targets were identified based on single-peptide assignment. Refine selection of the best representatives to be verified should reasonably performed. Literature-sourced OSCC biomarkers with previous evidence of presence in saliva, molecules identified with higher number of peptides, classifier discovery for the most cooperative proteins and commercially available antibodies or ELISA assays (as an alternative methodology) might be taken into consideration for the selection process. In summary, this study presented a quantitative-proteomic approach for discovery of OSCC candidate biomarkers in saliva of individuals from the Spanish population. We demonstrated a number of statistically significant proteins with the ability to distinguish OSCC patients from healthy controls and PVL based on their relative abundance profiles among which C3, APOA1, HP, HPX, SERPINA1, TYMP, TTR, ORM1, ORM2, CST4, ITIH1, S100A7, SERPINB4, SERPING1, protein AMBP, PFN1, LRG1, PLG, HSPA5, KNG1 could be associated with the risk of developing OSCC.

VI. *CONCLUSIONS*

1) IL-6, IL-8, TNF-α, HCC-1, and PF-4 showed significantly higher levels in OSCC than in OL and control saliva, being discriminately increased from early disease stages. These biomarkers may serve useful for definitive diagnosis, as well as for screening of patients at risk to develop oral cancer.

2) Correlation between HCC-1 and IL-6 levels and histologic features of OL lesions hints at the potential of these markers to indicate epithelial dysplasia. The considerable growth of TNF-α and IL-6 concentrations towards OSCC evolution, and IL-6 being distinctive in the presence of cervical adenopathy suggests their potential involvement in disease progression and severity.

3) IL-6 is suggested as an independent candidate biomarker with a capacity to discriminate between OSCC, OL, and control individuals.

4) Combined analytical strategies allowed the annotation of 90 putative salivary N-glycan structures. Differentially expressed fucosylated bi- and tri-antennary glycans between healthy, premalignant, and OSCC profiles suggest active N-glycosylation changes in the process of oral carcinogenesis and provide a reasonable platform for further investigation of their diagnostic potential.

5) Comparative proteome profiling revealed a list of differentially expressed proteins, characterized in OSCC, highlighting the importance of networks involving the immune system (Apolipoprotein A-I, Haptoglobin, Alpha-acid glycoprotein 1, Catalase, etc.), inhibition of enzymatic activities (Serpins family of protease inhibitors, Cathepsin D, protein AMBP and Cystatin-S) and cell adhesion (Cadherin-1, Vimentin, Vinculin, and Profilin-1) in this disease. Some newly found in saliva proteins (Calmodulin-like protein 5, Caspase-14, Filaggrin, High mobility group protein B2, Mucin 7, Tryptophanyl-tRNA synthetase, etc.) require further validation.

VII. *BIBLIOGRAPHY*

1. Döbróssy L. Epidemiology of head and neck cancer: Magnitude of the problem. Cancer and Metastasis Reviews. 2005.

2. Takahashi H, Fujita S, Tsuda N, Tezuka F, Okabe H. Intraoral Minor Salivary Gland Tumors: A Demographic and Histologic Study of 200 Cases. Tohoku J Exp Med. 1990;

3. Montero PH, Patel SG. Cancer of the Oral Cavity. Surgical Oncology Clinics of North America. 2015.

4. Gilyoma JM, Rambau PF, Masalu N, Kayange NM, Chalya PL. Head and neck cancers: A clinicopathological profile and management challenges in a resource-limited setting. BMC Res Notes. 2015;

5. Pindborg JJ, Reichart PA, Smith CJ, van der Waal I, Pindborg JJ, Reichart PA, et al. Histological Classification of Cancer and Precancer of the Oral Mucosa. In: Histological Typing of Cancer and Precancer of the Oral Mucosa. 1997.

6. Barnes L, Eveson JW, Reichart P, Sidransky D. World Health Organization Classification of Head and Neck Tumours. WHO Classif Tumour. 2005;

7. Bagan J, Sarrion G, Jimenez Y. Oral cancer: Clinical features. Oral Oncology. 2010.

8. Markopoulos AK. Current Aspects on Oral Squamous Cell Carcinoma. Open Dent J. 2012;

9. Alsahafi E, Begg K, Amelio I, Raulf N, Lucarelli P, Sauter T, et al. Clinical update on head and neck cancer: molecular biology and ongoing challenges. Cell Death Dis. 2019;

10. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;

11. Reibel J. Prognosis of oral pre-malignant lesions: Significance of clinical, histopathological, and molecular biological characteristics. Critical Reviews in Oral Biology and Medicine. 2003.

12. Rizzetto L, De Filippo C, Cavalieri D. Richness and diversity of mammalian fungal communities shape innate and adaptive immunity in health and disease. European Journal of Immunology. 2014.

13. Smith AJ, Jackson MS, Bagg J. The ecology of staphylococcus species in the oral cavity. Journal of Medical Microbiology. 2001.

14. Ling LJ, Hung SL, Tseng SC, Chen YT, Chi LY, Wu KM, et al. Association between betel quid chewin, periodontal status and periodontal pathogens. Oral Microbiol Immunol. 2001;

15. Börnigen D, Ren B, Pickard R, Li J, Ozer E, Hartmann EM, et al. Alterations in oral bacterial communities are associated with risk factors for oral and oropharyngeal cancer. Sci Rep. 2017;

16. Marttila E, Uittamo J, Rusanen P, Lindqvist C, Salaspuro M, Rautemaa R. Acetaldehyde production and microbial colonization in oral squamous cell carcinoma and oral lichenoid disease. Oral Surg Oral Med Oral Pathol Oral Radiol. 2013;

17. Rajeev R, Choudhary K, Panda S, Gandhi N. Role of bacteria in oral carcinogenesis. South Asian Journal of Cancer. 2012.

18. Banerjee S, Tian T, Wei Z, Peck KN, Shih N, Chalian AA, et al. Microbial Signatures Associated

with Oropharyngeal and Oral Squamous Cell Carcinomas. Sci Rep. 2017;

19. Ogden GR. Alcohol and oral cancer. In: Alcohol. 2005.

20. Leoncini E, Ricciardi W, Cadoni G, Arzani D, Petrelli L, Paludetti G, et al. Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer---systematic review and meta-analysis of trends by time and region. Head and Neck. 2014.

21. Kim SM. Human papilloma virus in oral cancer. J Korean Assoc Oral Maxillofac Surg. 2016;

22. Gonzalez-Moles MA, Gutierrez J, Rodriguez MJ, Ruiz-Avila I, Rodriguez-Archilla A. Epstein-Barr virus latent membrane protein-1 (LMP-1) expression in oral squamous cell carcinoma. Laryngoscope. 2002;

23. Hettmann A, Demcsák A, Decsi G, Bach, Pálinkó D, Rovó L, et al. Infectious agents associated with head and neck carcinomas. In: Advances in Experimental Medicine and Biology. 2016.

24. Warnakulasuriya S, Johnson NW, Van Der Waal I. Nomenclature and classification of potentially malignant disorders of the oral mucosa. J Oral Pathol Med. 2007;36(10):575–80.

25. van der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. Oral Oncology. 2009.

26. Wang YY, Tail YH, Wang WC, Chen CY, Kao YH, Chen YK, et al. Malignant transformation in 5071 southern Taiwanese patients with potentially malignant oral mucosal disorders. BMC Oral Health. 2014;

27. Bandyopadhyay A, Behura SS, Nishat R, Dash KC, Bhuyan L, Ramachandra S. Clinicopathological profile and malignant transformation in oral lichen planus: A retrospective study. Journal of International Society of Preventive and Community Dentistry. 2017.

28. Silverman S, Gorsky M. Proliferative verrucous leukoplakia: a follow-up study of 54 cases. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 1997;

29. Bagan J V., Jiménez-Soriano Y, Diaz-Fernandez JM, Murillo-Cortés J, Sanchis-Bielsa JM, Poveda-Roda R, et al. Malignant transformation of proliferative verrucous leukoplakia to oral squamous cell carcinoma: A series of 55 cases. Oral Oncol. 2011;

30. Amagasa T, Yamashiro M, Ishikawa H. Oral Leukoplakia Related to Malignant Transformation. Oral Sci Int. 2006;3(2):45–55.

31. Bagan J, Murillo-Cortes J, Poveda-Roda R, Leopoldo-Rodado M, Bagan L. Second primary tumors in proliferative verrucous leukoplakia: a series of 33 cases. Clin Oral Investig. 2019;

32. Bagan J, Scully C, Jimenez Y, Martorell M. Proliferative verrucous leukoplakia: A concise update. Oral Diseases. 2010.

33. Das BR, Nagpal JK. Understanding the biology of oral cancer. Medical Science Monitor. 2002.

34. Shirani S, Kargahi N, Razavi SM, Homayoni S. Epithelial dysplasia in oral cavity. Iranian Journal of Medical Sciences. 2014.

35. van der Meij EH, Mast H, van der Waal I. The possible premalignant character of oral lichen planus and oral lichenoid lesions: A prospective five-year follow-up study of 192 patients. Oral Oncol. 2007;

36. Ray JG, Ranganathan K, Chattopadhyay A. Malignant transformation of oral submucous fibrosis: overview of histopathological aspects. Oral Surgery, Oral Medicine, Oral Pathology

and Oral Radiology. 2016.

37. Dancyger A, Heard V, Huang B, Suley C, Tang D, Ariyawardana A. Malignant transformation of actinic cheilitis: A systematic review of observational studies. J Investig Clin Dent. 2018;

38. Scully C, Bagan J V. Oral squamous cell carcinoma: Overview of current understanding of aetiopathogenesis and clinical implications. Oral Diseases. 2009.

39. Lumerman H, Freedman P, Kerpel S. Oral epithelial dysplasia and the development of invasive squamous cell carcinoma. Oral Surgery, Oral Med Oral Pathol Oral Radiol. 1995;

40. Silverman S, Gorsky M, Lozada F. Oral leukoplakia and malignant transformation. A follow-up study of 257 patients. Cancer. 1984;

41. Saku T, Ohuchi T, Kohgo T, Shindo M, Katagiri M, Cheng J, et al. Carcinoma in-situ of the oral mucosa: Its pathological diagnostic concept based on the recognition of histological varieties proposed in the JSOP Oral CIS Catalog. J Oral Maxillofac Surgery, Med Pathol. 2014;

42. Arduino P, Bagan J, El-Naggar A, Carrozzo M. Urban legends series: Oral leukoplakia. Oral Diseases. 2013.

43. Awadallah M, Idle M, Patel K, Kademani D. Management update of potentially premalignant oral epithelial lesions. Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology. 2018.

44. Poh CF, Ng S, Berean KW, Williams PM, Rosin MP, Zhang L. Biopsy and histopathologic diagnosis of oral premalignant and malignant lesions. Journal of the Canadian Dental Association. 2008.

45. Macey R, Walsh T, Brocklehurst P, Kerr AR, Liu JLY, Lingen MW, et al. Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions. Cochrane Database of Systematic Reviews. 2015.

46. Brierley JD, Gospodarowicz MK, Wittekind C. TNM classification of malignant tumours - 8th edition. Union for International Cancer Control. 2017.

47. Rivera C. Essentials of oral cancer. International Journal of Clinical and Experimental Pathology. 2015.

48. Gao W, Guo C Bin. Factors Related to Delay in Diagnosis of Oral Squamous Cell Carcinoma. J Oral Maxillofac Surg. 2009;

49. Sujir N, Ahmed J, Pai K, Denny C, Shenoy N. Challenges in early diagnosis of oral cancer: Cases Series. Acta Stomatol Croat. 2019;

50. Mak MP, William WN. Targeting the epidermal growth factor receptor for head and neck cancer chemoprevention. Oral Oncol. 2014;

51. Epstein JB, Gorsky M, Wong FLW, Millner A. Topical bleomycin for the treatment of dysplastic oral leukoplakia. Cancer. 1998;

52. Foy JP, Bertolus C, William WN, Saintigny P. Oral premalignancy. The roles of early detection and chemoprevention. Otolaryngologic Clinics of North America. 2013.

53. Villa A, Woo S Bin. Leukoplakia—A Diagnostic and Management Algorithm. Journal of Oral and Maxillofacial Surgery. 2017.

54. Singh M, Krishanappa R, Bagewadi A, Keluskar V. Efficacy of oral lycopene in the treatment of oral leukoplakia. Oral Oncol. 2004;

55. Del Corso G, Gissi DB, Tarsitano A, Costabile E, Marchetti C, Montebugnoli L, et al. Laser evaporation versus laser excision of oral leukoplakia: A retrospective study with long-term follow-up. J Cranio-Maxillofacial Surg. 2015;

56. Van Der Hem PS, Nauta JM, Der Wal JEV, Roodenburg JLN. The results of CO 2 laser surgery in patients with oral leukoplakia: A 25 year follow up. Oral Oncol. 2005;

57. Asrani S, Reddy P, Dhirawani R, Jain S, Pathak S, Asati P. Cryosurgery: A simple tool to address oral lesions. Contemp Clin Dent. 2018;

58. Gaballah K, Costea DE, Hills A, Gollin SM, Harrison P, Partridge M. Tissue engineering of oral dysplasia. J Pathol. 2008;

59. Jakobsen C, Sørensen JA, Kassem M, Thygesen TH. Mesenchymal stem cells in oral reconstructive surgery: A systematic review of the literature. J Oral Rehabil. 2013;

60. Choi JH, Gimble JM, Lee K, Marra KG, Rubin JP, Yoo JJ, et al. Adipose tissue engineering for soft tissue regeneration. Tissue Engineering - Part B: Reviews. 2010.

61. Moharamzadeh K, Brook IM, Van Noort R, Scutt AM, Thornhill MH. Tissue-engineered oral mucosa: A review of the scientific literature. J Dent Res. 2007;

62. Lodi G, Franchini R, Warnakulasuriya S, Varoni EM, Sardella A, Kerr AR, et al. Interventions for treating oral leukoplakia to prevent oral cancer. Cochrane Database of Systematic Reviews. 2016.

63. Zelefsky MJ, Harrison LB, Fass DE, Armstrong JG, Shah JP, Strong EW. Postoperative radiation therapy for squamous cell carcinomas of the oral cavity and oropharynx: Impact of therapy on patients with positive surgical margins. Int J Radiat Oncol Biol Phys. 1993;

64. Mendenhall WM, Hinerman RW, Amdur RJ, Malyapa RS, Lansford CD, Werning JW, et al. Postoperative radiotherapy for squamous cell carcinoma of the head and neck. Clinical Medicine and Research. 2006.

65. Cooper JS, Pajak TF, Forastiere AA, Jacobs J, Campbell BH, Saxman SB, et al. Postoperative Concurrent Radiotherapy and Chemotherapy for High-Risk Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2004;

66. Bernier J, Domenge C, Ozsahin M, Matuszewska K, Lefèbvre JL, Greiner RH, et al. Postoperative Irradiation with or without Concomitant Chemotherapy for Locally Advanced Head and Neck Cancer. N Engl J Med. 2004;

67. Brizel DM, Albers ME, Fisher SR, Scher RL, Richtsmeier WJ, Hars V, et al. Hyperfractionated irradiation with or without concurrent chemotherapy for locally advanced head and neck cancer. N Engl J Med. 1998;

68. Robbins KT, Kumar P, Regine WF, Wong FSH, Weir AB, Flick P, et al. Efficacy of targeted supradose cisplatin and concomitant radiation therapy for advanced head and neck cancer: The Memphis experience. Int J Radiat Oncol Biol Phys. 1997;

69. Hamakawa H, Nakashiro KI, Sumida T, Shintani S, Myers JN, Takes RP, et al. Basic evidence of molecular targeted therapy for oral cancer and salivary gland cancer. Head and Neck. 2008.

70. Sim F, Leidner R, Bell RB. Immunotherapy for Head and Neck Cancer. Oral and Maxillofacial Surgery Clinics of North America. 2019.

71. Prasad V, Kaestner V. Nivolumab and pembrolizumab: Monoclonal antibodies against

programmed cell death-1 (PD-1) that are interchangeable. Seminars in Oncology. 2017.

72. Ferris RL, Blumenschein G, Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab vs investigator's choice in recurrent or metastatic squamous cell carcinoma of the head and neck: 2-year long-term survival update of CheckMate 141 with analyses by tumor PD-L1 expression. Oral Oncol. 2018;

73. Güneri P, Epstein JB. Late stage diagnosis of oral cancer: Components and possible solutions. Oral Oncology. 2014.

74. Kowalski LP, Carvalho AL. Natural history of untreated head and neck cancer. Eur J Cancer. 2000;

75. de Cecco L, Nicolau M, Giannoccaro M, Daidone MG, Bossi P, Locati L, et al. Head and neck cancer subtypes with biological and clinical relevance: Meta-analysis of gene-expression data. Oncotarget. 2015;

76. Leõn X, Martínez V, Lõpez M, García J, Venegas MDP, Esteller E, et al. Second, third, and fourth head and neck tumors. A progressive decrease in survival. Head Neck. 2012;

77. Choi S, Myers JN. Molecular pathogenesis of oral squamous cell carcinoma: Implications for therapy. J Dent Res. 2008;87(1):14–32.

78. Sidransky D. Molecular genetics of head and neck cancer. Current Opinion in Oncology. 1995.

79. Califano J, Van Der Riet P, Westra W, Nawroz H, Clayman G, Piantadosi S, et al. Genetic progression model for head and neck cancer: Implications for field cancerization. Cancer Res. 1996;

80. Garnis C, Baldwin C, Zhang L, Rosin MP, Lam WL. Use of Complete Coverage Array Comparative Genomic Hybridization to Define Copy Number Alterations on Chromosome 3p in Oral Squamous Cell Carcinomas. Cancer Res. 2003;

81. Masayesva BG, Ha P, Garrett-Mayer E, Pilkington T, Mao R, Pevsner J, et al. Gene expression alterations over large chromosomal regions in cancers include multiple genes unrelated to malignant progression. Proc Natl Acad Sci U S A. 2004;

82. Rousseau A, Lim MS, Lin Z, Jordan RCK. Frequent cyclin D1 gene amplification and protein overexpression in oral epithelial dysplasias. Oral Oncol. 2001;

83. Slaughter DP, Southwick HW, Smejkal W. "Field cancerization" in oral stratified squamous epithelium. Clinical implications of multicentric origin. Cancer. 1953;

84. Feller LL, Khammissa RRAG, Kramer BB, Lemmer JJ. Oral squamous cell carcinoma in relation to field precancerisation: Pathobiology. Cancer Cell International. 2013.

85. González-Moles MA, Scully C, Ruiz-Ávila I, Plaza-Campillo JJ. The cancer stem cell hypothesis applied to oral carcinoma. Oral Oncology. 2013.

86. Kumar SM, Liu S, Lu H, Zhang H, Zhang PJ, Gimotty PA, et al. Acquired cancer stem cell phenotypes through Oct4-mediated dedifferentiation. Oncogene. 2012;

87. Squier C, Brogden KA. Human Oral Mucosa: Development, Structure and Function. Human Oral Mucosa: Development, Structure and Function. 2013.

88. Simple M, Suresh A, Das D, Kuriakose MA. Cancer stem cells and field cancerization of Oral squamous cell carcinoma. Oral Oncology. 2015.

89. Bedi GC, Westra WH, Gabrielson E, Koch W, Sidransky D. Multiple head and neck tumors: Evidence for a common clonal origin. Cancer Res. 1996;

90. Ha PK, Benoit NE, Yochem R, Sciubba J, Zahurak M, Sidransky D, et al. A transcriptional progression model for head and neck cancer. Clin Cancer Res. 2003;

91. Braakhuis BJM, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of slaughter's concept of field cancerization: Evidence and clinical implications. Cancer Research. 2003.

92. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000.

93. Singh B, Carpenter G, Coffey RJ. EGF receptor ligands: recent advances. F1000Research. 2016;

94. Rogers SJ, Harrington KJ, Rhys-Evans P, O-Charoenrat P, Eccles SA. Biological significance of c-erbB family oncogenes in head and neck cancer. Cancer and Metastasis Reviews. 2005.

95. Kalyankrishna S, Grandis JR. Epidermal growth factor receptor biology in head and neck cancer. Journal of Clinical Oncology. 2006.

96. Chau NG, Perez-Ordonez B, Zhang K, Pham NA, Ho J, Zhang T, et al. The association between EGFR variant III, HPV, p16, c-MET, EGFR gene copy number and response to EGFR inhibitors in patients with recurrent or metastatic squamous cell carcinoma of the head and neck. Head Neck Oncol. 2011;

97. Ali SMA, Mirza Y. Overexpression of EGFR, COX2 and p53 in oral squamous cell carcinoma patients of Pakistan and correlation with prognosis. Ann Oncol. 2019;

98. Solomon B, Young RJ, Rischin D. Head and neck squamous cell carcinoma: Genomics and emerging biomarkers for immunomodulatory cancer treatments. Seminars in Cancer Biology. 2018.

99. Hong WK, Ro JY, Hittelman WN. Dysregulation of Epidermal Growth Factor Receptor Expression in Premalignant Lesions during Head and Neck Tumorigenesis. Cancer Res. 1994;

100. Costa V, Kowalski LP, Coutinho-Camillo CM, Begnami MD, Calsavara VF, Neves JI, et al. EGFR amplification and expression in oral squamous cell carcinoma in young adults. Int J Oral Maxillofac Surg. 2018;

101. Ma PC, Maulik G, Christensen J, Salgia R. c-Met: Structure, functions and potential for therapeutic inhibition. Cancer and Metastasis Reviews. 2003.

102. Cho YA, Kim EK, Heo SJ, Cho BC, Kim HR, Chung JM, et al. Alteration status and prognostic value of MET in head and neck squamous cell carcinoma. J Cancer. 2016;

103. Szturz P, Raymond E, Abitbol C, Albert S, de Gramont A, Faivre S. Understanding c-MET signalling in squamous cell carcinoma of the head & neck. Critical Reviews in Oncology/Hematology. 2017.

104. Morello S, Olivero M, Aimetti M, Bernardi M, Berrone S, Di Renzo MF, et al. MET receptor is overexpressed but not mutated in oral squamous cell carcinomas. J Cell Physiol. 2001;

105. Leemans CR, Braakhuis BJM, Brakenhoff RH. The molecular biology of head and neck cancer. Nature Reviews Cancer. 2011.

106. Peng Q, Deng Z, Pan H, Gu L, Liu O, Tang Z. Mitogen-activated protein kinase signaling pathway in oral cancer (Review). Oncology Letters. 2018.

107. Das N, Majumder J, Dasgupta UB. ras Gene mutations in oral cancer in eastern India. Oral Oncol. 2000;

108. S.E. C, P. B, N.W. J, P.R. M, F. M, B. Y, et al. Ras mutations in United Kingdom examples of oral malignancies are infrequent. Int J Cancer. 1991;

109. Anderson JA, Irish JC, Mclachlin CM, Ngan BY. H-ras Oncogene Mutation and Human Papillomavirus Infection in Oral Carcinomas. Arch Otolaryngol Neck Surg. 1994;

110. De Carvalho TG, De Carvalho AC, Maia DCC, Ogawa JK, Carvalho AL, Vettore AL. Search for mutations in signaling pathways in head and neck squamous cell carcinoma. Oncol Rep. 2013;

111. Rothenberg SM, Ellisen LW. The molecular pathogenesis of head and neck squamous cell carcinoma. Journal of Clinical Investigation. 2012.

112. Yang K, Hitomi M, Stacey DW. Variations in cyclin D1 levels through the cell cycle determine the proliferative fate of a cell. Cell Division. 2006.

113. Ortiz AB, Garcia D, Vicente Y, Palka M, Bellas C, Martin P. Prognostic significance of cyclin D1 protein expression and gene amplification in invasive breast carcinoma. PLoS One. 2017;

114. Miyamoto R, Uzawa N, Nagaoka S, Hirata Y, Amagasa T. Prognostic significance of cyclin D1 amplification and overexpression in oral squamous cell carcinomas. Oral Oncol. 2003;

115. Lam KY, Ng IOL, Yuen APW, Kwong DLW, Wei W. Cyclin D1 expression in oral squamous cell carcinomas: Clinicopathological relevance and correlation with p53 expression. J Oral Pathol Med. 2000;

116. Mineta H, Miura K, Takebayashi S, Ueda Y, Misawa K, Harada H, et al. Cyclin D1 overexpression correlates with poor prognosis in patients with tongue squamous cell carcinoma. Oral Oncol. 2000;

117. Yu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: A leading role for STAT3. Nature Reviews Cancer. 2009.

118. Bowman T, Garcia R, Turkson J, Jove R. STATs in oncogenesis. Oncogene. 2000.

119. Geiger JL, Grandis JR, Bauman JE. The STAT3 pathway as a therapeutic target in head and neck cancer: Barriers and innovations. Oral Oncology. 2016.

120. Grandis JR, Drenning SD, Zeng Q, Watkins SC, Melhem MF, Endo S, et al. Constitutive activation of stat3 signaling abrogates apoptosis in squamous cell carcinogenesis in vivo. Proc Natl Acad Sci U S A. 2000;

121. Masuda M, Suzui M, Yasumatu R, Nakashima T, Kuratomi Y, Azuma K, et al. Constitutive activation of signal transducers and activators of transcription 3 correlates with cyclin D1 overexpression and may provide a novel prognostic marker in head and neck squamous cell carcinoma. Cancer Res. 2002;

122. Mishra A, Bharti AC, Varghese P, Saluja D, Das BC. Differential expression and activation of NF-κB family proteins during oral carcinogenesis: Role of high risk human papillomavirus infection. Int J Cancer. 2006;

123. Kamperos G, Nikitakis N, Sfakianou A, Avgoustidis D, Sklavounou-Andrikopoulou A. Expression of NF-κB and IL-6 in oral precancerous and cancerous lesions: An immunohistochemical study. Med Oral Patol Oral Cir Bucal. 2016;

124. Lopez-Bergami P, Lau E, Ronai Z. Emerging roles of ATF2 and the dynamic AP1 network in cancer. Nature Reviews Cancer. 2010.

125. Turatti E, da Costa Neves A, de Magalhães MHCG, de Sousa SOM. Assessment of c-Jun, c-Fos and cyclin D1 in premalignant and malignant oral lesions. J Oral Sci. 2005;

126. Robinson CM, Prime SS, Huntley S, Stone AM, Davies M, Eveson JW, et al. Overexpression of JunB in undifferentiated malignant rat oral keratinocytes enhances the malignant phenotype in vitro without altering cellular differentiation. Int J Cancer. 2001;

127. Evans HJ, Prosser J. Tumour-supresssr genes: Cardinal factors in inherited predisposition to human cancers. In: Environmental Health Perspectives. 1992.

128. Pande P, Mathur M, Shukla NK, Ralhan R. pRB and p16 protein alterations in human oral tumorigenesis. Oral Oncol. 1998;

129. Koontongkaew S, Chareonkitkajorn L, Chanvitan A, Leelakriangsak M, Amornphimoltham P. Alterations of p53, pRb, cyclin D1 and cdk4 in human oral and pharyngeal squamous cell carcinomas. Oral Oncol. 2000;

130. Sullivan KD, Galbraith MD, Andrysik Z, Espinosa JM. Mechanisms of transcriptional regulation by p53. Cell Death and Differentiation. 2018.

131. Aubrey BJ, Kelly GL, Janic A, Herold MJ, Strasser A. How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? Cell Death and Differentiation. 2018.

132. Zhou G, Liu Z, Myers JN. TP53 Mutations in Head and Neck Squamous Cell Carcinoma and Their Impact on Disease Progression and Treatment Response. J Cell Biochem. 2016;

133. Castellsagué X, Alemany L, Quer M, Halec G, Quirós B, Tous S, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. J Natl Cancer Inst. 2016;

134. Ghanghoria S, Ghanghoria A, Shukla A. p53 Expression in Oral cancer: A study of 50 cases. J Pathol Nepal. 2015;

135. Abbas T, Dutta A. P21 in cancer: Intricate networks and multiple activities. Nature Reviews Cancer. 2009.

136. Agarwal S, Mathur M, Shukla NK, Ralhan R. Expression of cyclin dependent kinase inhibitor p21(waf1/cip1) in premalignant and malignant oral lesions: Relationship with p53 status. Oral Oncol. 1998;

137. Ortega S, Malumbres M, Barbacid M. Cyclin D-dependent kinases, INK4 inhibitors and cancer. Biochimica et Biophysica Acta - Reviews on Cancer. 2002.

138. Thambiah LJ, Bindushree R V., Anjum A, Pugazhendi SK, Babu L, Nair RP. Evaluating the expression of p16 and p27 in oral epithelial dysplasias and oral squamous cell carcinoma: A diagnostic marker for carcinogenesis. J Oral Maxillofac Pathol. 2018;

139. Agarwal A, Kamboj M, Shreedhar B. “Expression of p16 in oral leukoplakia and oral squamous cell carcinoma and correlation of its expression with individual atypical features.” J Oral Biol Craniofacial Res. 2019;

140. Bova RJ, Quinn DI, Nankervis JS, Cole IE, Sheridan BF, Jensen MJ, et al. Cyclin D1 and p16(INK4A) expression predict reduced survival in carcinoma of the anterior tongue. Clin Cancer Res. 1999;

141. Artavanis-Tsakonas S, Rand MD, Lake RJ. Notch signaling: Cell fate control and signal integration in development. Science. 1999.

142. Dotto GP. Notch tumor suppressor function. Oncogene. 2008.

143. Agrawal N, Frederick MJ, Pickering CR, Bettegowda C, Chang K, Li RJ, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science (80-). 2011;

144. Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. Nature Reviews Cancer. 2002.

145. Birchmeier W, Behrens J. Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness. BBA - Reviews on Cancer. 1994.

146. Ukpo OC, Thorstad WL, Zhang Q, Lewis JS. Lack of Association of Cadherin Expression and Histopathologic Type, Metastasis, or Patient Outcome in Oropharyngeal Squamous Cell Carcinoma: A Tissue Microarray Study. Head Neck Pathol. 2012;

147. Lim SC, Zhang S, Ishii G, Endoh Y, Kodama K, Miyamoto S, et al. Predictive Markers for Late Cervical Metastasis in Stage I and II Invasive Squamous Cell Carcinoma of the Oral Tongue. Clin Cancer Res. 2004;

148. Fan CC, Wang TY, Cheng YA, Jiang SS, Cheng CW, Lee AYL, et al. Expression of E-cadherin, Twist, and p53 and their prognostic value in patients with oral squamous cell carcinoma. J Cancer Res Clin Oncol. 2013;

149. Diniz-Freitas M, García-Caballero T, Antúnez-López J, Gándara-Rey JM, García-García A. Reduced E-cadherin expression is an indicator of unfavourable prognosis in oral squamous cell carcinoma. Oral Oncol. 2006;

150. von Zeidler SV, de Souza Botelho T, Mendonça EF, Batista AC. E-cadherin as a potential biomarker of malignant transformation in oral leukoplakia: A retrospective cohort study. BMC Cancer. 2014;

151. Thomas GJ, Speight PM. Cell adhesion molecules and oral cancer. Critical Reviews in Oral Biology and Medicine. 2001.

152. Thomas GJ, Jones J, Speight PM. Integrins and oral cancer. Oral Oncology. 1997.

153. Hamidi S, Salo T, Kainulainen T, Epstein J, Lerner K, Larjava H. Expression of α(v)β6 integrin in oral leukeplakia. Br J Cancer. 2000;

154. Thomas GJ, Nyström ML, Marshall JF. αvβ6 integrin in wound healing and cancer of the oral cavity. Journal of Oral Pathology and Medicine. 2006.

155. Xue H, Atakilit A, Zhu W, Li X, Ramos DM, Pytela R. Role of the αvβ6 integrin in human oral squamous cell carcinoma growth in vivo and in vitro. Biochem Biophys Res Commun. 2001;

156. Wong TS, Gao W, Li ZH. Matrix metalloproteinase family as molecular biomarkers in oral squamous cell carcinoma. In: Biomarkers in Disease: Methods, Discoveries and Applications: Biomarkers in Cancer. 2015.

157. Kawamata H, Uchida D, Hamano H, Kimura-Yanagawa T, Nakashiro KI, Hino S, et al. Active-MMP2 in cancer cell nests of oral cancer patients: Correlation with lymph node metastasis. Int J Oncol. 1998;

158. De Vicente JC, Fresno MF, Villalain L, Vega JA, Hernández Vallejo G. Expression and clinical

significance of matrix metalloproteinase-2 and matrix metalloproteinase-9 in oral squamous cell carcinoma. Oral Oncol. 2005;

159. Zhang JM, An J. Cytokines, inflammation, and pain. International Anesthesiology Clinics. 2007.

160. Zlotnik A, Yoshie O. The Chemokine Superfamily Revisited. Immunity. 2012.

161. Hasturk H, Kantarci A, Van Dyke TE. Oral inflammatory diseases and systemic inflammation: Role of the macrophage. Frontiers in Immunology. 2012.

162. Gregory AD, Houghton AMG. Tumor-associated neutrophils: New targets for cancer therapy. Cancer Research. 2011.

163. Zamarron BF, Chen W. Dual roles of immune cells and their factors in cancer development and progression. Int J Biol Sci. 2011;

164. S.P. H, C.C. H. Inflammation and cancer: An ancient link with novel potentials. International Journal of Cancer. 2007.

165. Grivennikov SI, Greten FR, Karin M. Immunity, Inflammation, and Cancer. Cell. 2010.

166. Lin WW, Karin M. A cytokine-mediated link between innate immunity, inflammation, and cancer. Journal of Clinical Investigation. 2007.

167. Grivennikov SI, Karin M. Inflammatory cytokines in cancer: Tumour necrosis factor and interleukin 6 take the stage. In: Annals of the Rheumatic Diseases. 2011.

168. Lawrence T. The nuclear factor NF-kappaB pathway in inflammation. Cold Spring Harbor perspectives in biology. 2009.

169. Chen Z, Malhotra PS, Thomas GR, Ondrey FG, Duffey DC, Smith CW, et al. Expression of proinflammatory and proangiogenic cytokines in patients with head and neck cancer. Clin Cancer Res. 1999;

170. Rhodus NL, Ho V, Miller CS, Myers S, Ondrey F. NF-κB dependent cytokine levels in saliva of patients with oral preneoplastic lesions and oral squamous cell carcinoma. Cancer Detect Prev. 2005;

171. Dikova VR, Principe S, Bagan J V. Salivary inflammatory proteins in patients with oral potentially malignant disorders. Journal of Clinical and Experimental Dentistry. 2019.

172. Brailo V, Vucicevic-Boras V, Lukac J, Biocina-Lukenda D, Zilic-Alajbeg I, Milenovic A, et al. Salivary and serum interleukin 1 beta, interleukin 6 and tumor necrosis factor alpha in patients with leukoplakia and oral cancer. Med Oral Patol Oral Cir Bucal. 2012;

173. Kaur J, Jacobs R. Proinflammatory cytokine levels in oral lichen planus, oral leukoplakia, and oral submucous fibrosis. J Korean Assoc Oral Maxillofac Surg. 2015;

174. Nakayama H, Ikebe T, Beppu M, Shirasuna K. High expression levels of nuclear factor κB, IκB kinase α and akt kinase in squamous cell carcinoma of the oral cavity. Cancer. 2001;

175. Korostoff A, Reder L, Masood R, Sinha UK. The role of salivary cytokine biomarkers in tongue cancer invasion and mortality. Oral Oncol. 2011;

176. Bossi P, Bergamini C, Miceli R, Cova A, Orlandi E, Resteghini C, et al. Salivary Cytokine Levels and Oral Mucositis in Head and Neck Cancer Patients Treated With Chemotherapy and Radiation Therapy. Int J Radiat Oncol Biol Phys. 2016;

177. Ellenberg SS, Hamilton JM. Surrogate endpoints in clinical trials: Cancer. Stat Med. 1989;

178. Silberring J, Ciborowski P. Biomarker discovery and clinical proteomics. TrAC - Trends in Analytical Chemistry. 2010.

179. Mayeux R. Biomarkers: Potential Uses and Limitations. NeuroRx. 2004;

180. Moncada V, Srivastava S. Biomarkers in oncology research and treatment: Early detection research network: A collaborative approach. Biomark Med. 2008;

181. Rusling JF, Kumar C V., Gutkind JS, Patel V. Measurement of biomarker proteins for point-of-care early detection and monitoring of cancer. Analyst. 2010;

182. Liu ET. Systems biology, integrative biology, predictive biology. Cell. 2005.

183. Nagaraj NS. Evolving "omics" technologies for diagnostics of head and neck cancer. Briefings Funct Genomics Proteomics. 2009;

184. Wu RQ, Zhao XF, Wang ZY, Zhou M, Chen QM. Novel molecular events in oral carcinogenesis via integrative approaches. Journal of Dental Research. 2011.

185. Gao J, Garulacan LA, Storm SM, Opiteck GJ, Dubaquie Y, Hefta SA, et al. Biomarker discovery in biological fluids. Methods. 2005;

186. Kaczor-Urbanowicz KE, Martin Carreras-Presas C, Aro K, Tu M, Garcia-Godoy F, Wong DTW. Saliva diagnostics – Current views and directions. Experimental Biology and Medicine. 2017.

187. Yoshizawa JM, Schafer CA, Schafer JJ, Farrell JJ, Paster BJ, Wong DTW. Salivary biomarkers: Toward future clinical and diagnostic utilities. Clinical Microbiology Reviews. 2013.

188. 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;

189. Javaid MA, Ahmed AS, Durand R, Tran SD. Saliva as a diagnostic tool for oral and systemic diseases. Journal of Oral Biology and Craniofacial Research. 2016.

190. Sinevici N, O'sullivan J. Oral cancer: Deregulated molecular events and their use as biomarkers. Oral Oncology. 2016.

191. Wu JY, Yi C, Chung HR, Wang DJ, Chang WC, Lee SY, et al. Potential biomarkers in saliva for oral squamous cell carcinoma. Vol. 46, Oral Oncology. 2010. p. 226–31.

192. Malik UU, Zarina S, Pennington SR. Oral squamous cell carcinoma: Key clinical questions, biomarker discovery, and the role of proteomics. Archives of Oral Biology. 2016.

193. Sepiashvili L, Hui A, Ignatchenko V, Shi W, Su S, Xu W, et al. Potentially novel candidate biomarkers for head and neck squamous cell carcinoma identified using an integrated cell line-based discovery strategy. Mol Cell Proteomics. 2012;

194. Kawahara R, Bollinger JG, Rivera C, Ribeiro ACP, Brandão TB, Leme AFP, et al. A targeted proteomic strategy for the measurement of oral cancer candidate biomarkers in human saliva. Proteomics. 2016;

195. Oppenheim FG, Salih E, Siqueira WL, Zhang W, Helmerhorst EJ. Salivary proteome and its genetic polymorphisms. In: Annals of the New York Academy of Sciences. 2007.

196. Lo WY, Tsai MH, Tsai Y, Hua CH, Tsai FJ, Huang SY, et al. Identification of over-expressed proteins in oral squamous cell carcinoma (OSCC) patients by clinical proteomic analysis. Clin Chim Acta. 2007;

197. Katakura A, Kamiyama I, Takano N, Shibahara T, Muramatsu T, Ishihara K, et al. Comparison of salivary cytokine levels in oral cancer patients and healthy subjects. Bull Tokyo Dent Coll. 2007;

198. Franzmann EJ, Reategui EP, Pedroso F, Pernas FG, Karakullukcu BM, Carraway KL, et al. Soluble CD44 is a potential marker for the early detection of head and neck cancer. Cancer Epidemiol Biomarkers Prev. 2007;

199. Nagler R, Bahar G, Shpitzer T, Feinmesser R. Concomitant analysis of salivary tumor markers - A new diagnostic tool for oral cancer. Clin Cancer Res. 2006;

200. Dowling P, Wormald R, Meleady P, Henry M, Curran A, Clynes M. Analysis of the saliva proteome from patients with head and neck squamous cell carcinoma reveals differences in abundance levels of proteins associated with tumour progression and metastasis. J Proteomics. 2008;

201. Pickering V, Jordan RCK, Schmidt BL. Elevated salivary endothelin levels in oral cancer patients-A pilot study. Oral Oncol. 2007;

202. Principe S, Hui ABY, Bruce J, Sinha A, Liu FF, Kislinger T. Tumor-derived exosomes and microvesicles in head and neck cancer: Implications for tumor biology and biomarker discovery. Proteomics. 2013.

203. Gallo C, Ciavarella D, Santarelli A, Ranieri E, Colella G, Muzio L Lo, et al. Potential salivary proteomic markers of oral squamous cell carcinoma. Cancer Genomics and Proteomics. 2016;

204. Mechref Y, Hu Y, Garcia A, Hussein A. Identifying cancer biomarkers by mass spectrometry-based glycomics. Electrophoresis. 2012.

205. Jensen PH, Karlsson NG, Kolarich D, Packer NH. Structural analysis of N- and O-glycans released from glycoproteins. Nat Protoc. 2012;

206. Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al. Essentials of glycobiology, third edition. Cold Spring Harbor Laboratory Press. 2017.

207. Adamczyk B, Tharmalingam T, Rudd PM. Glycans as cancer biomarkers. Biochimica et Biophysica Acta - General Subjects. 2012.

208. Guu SY, Lin TH, Chang SC, Wang RJ, Hung LY, Fang PJ, et al. Serum N-glycome characterization and anti-carbohydrate antibody profiling in oral squamous cell carcinoma patients. PLoS One. 2017;

209. Lin MC, Huang MJ, Liu CH, Yang TL, Huang MC. GALNT2 enhances migration and invasion of oral squamous cell carcinoma by regulating EGFR glycosylation and activity. Oral Oncol. 2014;

210. Sinevici N, Mittermayr S, Davey GP, Bones J, O'Sullivan J. Salivary N-glycosylation as a biomarker of oral cancer: A pilot study. Glycobiology. 2019;

211. van der Waal I. Oral leukoplakia, the ongoing discussion on definition and terminology. Med Oral Patol Oral Cir Bucal. 2015;

212. NAVAZESH M. Methods for Collecting Saliva. Ann N Y Acad Sci. 1993;

213. Helmerhorst EJ, Oppenheim FG. Saliva: A dynamic proteome. Journal of Dental Research. 2007.

214. Perlee LT, Christiansen J, Dondero R, Grimwade B, Lejnine S, Mullenix M, et al. Development and standardization of multiplexed antibody microarrays for use in quantitative proteomics. Proteome Sci. 2004;

215. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976;

216. Carillo S, Mittermayr S, Farrell A, Albrecht S, Bones J. Glycosylation analysis of therapeutic glycoproteins produced in CHO cells. In: Methods in Molecular Biology. 2017.

217. Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. J Proteome Res. 2008;

218. Zor T, Selinger Z. Linearization of the Bradford protein assay increases its sensitivity: Theoretical and experimental studies. Anal Biochem. 1996;

219. Shevchenko A, Jensen ON, Podtelejnikov A V., Sagliocco F, Wilm M, Vorm O, et al. Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels. In: Proceedings of the National Academy of Sciences of the United States of America. 1996.

220. Shilov I V., Seymourt SL, Patel AA, Loboda A, Tang WH, Keating SP, et al. The paragon algorithm, a next-generation search engine that uses sequence temperature values sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. Mol Cell Proteomics. 2007;

221. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. BMC Genet. 2010;

222. Hastie T, Tibshirani R, Friedman J. Elements of Statistical Learning 2nd ed. Elements. 2009.

223. James G, Witten D, Hastie T, Tibshirani R. Introduction to Statistical Learning with Applications in R. Springer New York. 2013.

224. Key M. A tutorial in displaying mass spectrometry-based proteomic data using heat maps. BMC bioinformatics. 2012.

225. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;

226. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;

227. Wong DT. Towards a simple, saliva-based test for the detection of oral cancer. “Oral fluid (saliva), which is the mirror of the body, is a perfect medium to be explored for health and disease surveillance.” Expert Review of Molecular Diagnostics. 2006.

228. Almadori G, Bussu F, Galli J, Limongelli A, Persichilli S, Zappacosta B, et al. Salivary glutathione and uric acid levels in patients with head and neck squamous cell carcinoma. Head Neck. 2007;

229. Osman TA, Costea DE, Johannessen AC. The use of salivary cytokines as a screening tool for oral squamous cell carcinoma: A review of the literature. Journal of Oral and Maxillofacial Pathology. 2012.

230. Sahibzada HA, Khurshid Z, Khan RS, Naseem M, Siddique KM, Mali M, et al. Salivary IL-8, IL-

6 and TNF-α as Potential Diagnostic Biomarkers for Oral Cancer. Diagnostics. 2017;7(2):21.

231. Prasad G, McCullough M. Chemokines and cytokines as salivary biomarkers for the early diagnosis of oral Cancer. Int J Dent. 2013;

232. Babiuch K, Kuśnierz-Cabala B, Kęsek B, Okoń K, Darczuk D, Chomyszyn-Gajewska M. Evaluation of Proinflammatory, NF-kappaB Dependent Cytokines: IL-1α, IL-6, IL-8, and TNF-α in Tissue Specimens and Saliva of Patients with Oral Squamous Cell Carcinoma and Oral Potentially Malignant Disorders. J Clin Med. 2020;

233. Juretić M, Cerović R, Belušić-Gobić M, Pršo IB, Kqiku L, Špalj S, et al. Salivary levels of TNF-α and IL-6 in patients with oral premalignant and malignant lesions. Folia Biol (Czech Republic). 2013;

234. Dineshkumar T, Ashwini BK, Rameshkumar A, Rajashree P, Ramya R, Rajkumar K. Salivary and serum interleukin-6 levels in oral premalignant disorders and squamous cell Carcinoma: Diagnostic value and clinicopathologic correlations. Asian Pacific J Cancer Prev. 2016;

235. Bagan L, Sáez GT, Tormos MC, Labaig-Rueda C, Murillo-Cortes J, Bagan J V. Salivary and serum interleukin-6 levels in proliferative verrucous leukoplakia. Clin Oral Investig. 2016;

236. Shurin G V., Ferris R, Tourkova IL, Perez L, Lokshin A, Balkir L, et al. Loss of New Chemokine CXCL14 in Tumor Tissue Is Associated with Low Infiltration by Dendritic Cells (DC), while Restoration of Human CXCL14 Expression in Tumor Cells Causes Attraction of DC Both In Vitro and In Vivo. J Immunol. 2005;

237. Sharma M, Bairy I, Pai K, Satyamoorthy K, Prasad S, Berkovitz B, et al. Salivary IL-6 levels in oral leukoplakia with dysplasia and its clinical relevance to tobacco habits and periodontitis. Clin Oral Investig. 2011;

238. Pries R, Wollenberg B. Cytokines in head and neck cancer. Cytokine and Growth Factor Reviews. 2006.

239. Schapher M, Wendler O, Gröschl M. Salivary cytokines in cell proliferation and cancer. Clinica Chimica Acta. 2011.

240. Leibowitz MS, Nayak J V., Ferris RL. Head and neck cancer immunotherapy: Clinical evaluation. Current Oncology Reports. 2008.

241. Gales D, Clark C, Manne U, Samuel T. The Chemokine CXCL8 in Carcinogenesis and Drug Response. ISRN Oncol. 2013;

242. Krishnan R, Thayalan DK, Padmanaban R, Ramadas R, Annasamy RK, Anandan N. Association of serum and salivary tumor necrosis factor-α with histological grading in oral cancer and its role in differentiating premalignant and malignant oral disease. Asian Pacific J Cancer Prev. 2014;

243. St. John MAR, Li Y, Zhou X, Denny P, Ho CM, Montemagno C, et al. Interleukin 6 and interleukin 8 as potential biomarkers for oral cavity and oropharyngeal squamous cell carcinoma. Arch Otolaryngol - Head Neck Surg. 2004;

244. Sahebjamee M, Eslami M, Atarbashimoghadam F, Sarafnejad A. Salivary concentration of TNFα, IL1α, IL6, and IL8 in oral squamous cell carcinoma. Med Oral Patol Oral Cir Bucal. 2008;

245. Lee LT, Wong YK, Hsiao HY, Wang YW, Chan MY, Chang KW. Evaluation of saliva and plasma cytokine biomarkers in patients with oral squamous cell carcinoma. Int J Oral Maxillofac Surg. 2018;

246. Mukaida N, Sasaki SI, Baba T. Chemokines in cancer development and progression and their potential as targeting molecules for cancer treatment. Mediators of Inflammation. 2014.

247. Liu KYP, Lu XJD, Zhu YS, Le N, Kim H, Poh CF. Plasma-derived inflammatory proteins predict oral squamous cell carcinoma. Front Oncol. 2018;

248. Yeudall WA, Miyazaki H. Chemokines and squamous cancer of the head and neck: Targets for therapeutic intervention? Expert Review of Anticancer Therapy. 2007.

249. Michiels K, Schutyser E, Conings R, Lenaerts JP, Put W, Nuyts S, et al. Carcinoma cell-derived chemokines and their presence in oral fluid. Eur J Oral Sci. 2009;

250. Tsou CL, Peters W, Si Y, Slaymaker S, Aslanian AM, Weisberg SP, et al. Critical roles for CCR2 and MCP-3 in monocyte mobilization from bone marrow and recruitment to inflammatory sites. J Clin Invest. 2007;

251. Ferreira FO, Ribeiro FLL, Batista AC, Leles CR, De Cássia Gonçalves Alencar R, Silva TA. Association of CCL2 with lymph node metastasis and macrophage infiltration in oral cavity and lip squamous cell carcinoma. Tumor Biol. 2008;

252. Ohta M, Kitadai Y, Tanaka S, Yoshihara M, Yasui W, Mukaida N, et al. Monocyte chemoattractant protein-1 expression correlates with macrophage infiltration and tumor vascularity in human esophageal squamous cell carcinomas. Int J Cancer. 2002;

253. Cai Y, Ling Y, Huang L, Huang H, Chen X, Xiao Y, et al. C-C motif chemokine 14 as a novel potential biomarker for predicting the prognosis of epithelial ovarian cancer. Oncol Lett. 2020;

254. Feng L, Houck JR, Lohavanichbutr P, Chen C. Transcriptome analysis reveals differentially expressed lncRNAs between oral squamous cell carcinoma and healthy oral mucosa. Oncotarget. 2017;

255. Westrich JA, Vermeer DW, Colbert PL, Spanos WC, Pyeon D. The multifarious roles of the chemokine CXCL14 in cancer progression and immune responses. Mol Carcinog. 2020;

256. Hata R-I. A New Strategy to Find Targets for Anticancer Therapy: Chemokine CXCL14/BRAK Is a Multifunctional Tumor Suppressor for Head and Neck Squamous Cell Carcinoma. ISRN Otolaryngol. 2012;

257. Ozawa S, Kato Y, Komori R, Maehata Y, Kubota E, Hata RI. BRAK/CXCL14 expression suppresses tumor growth in vivo in human oral carcinoma cells. Biochem Biophys Res Commun. 2006;

258. Frederick MJ, Henderson Y, Xu X, Deavers MT, Sahin AA, Wu H, et al. In vivo expression of the novel CXC chemokine BRAK in normal and cancerous human tissue. Am J Pathol. 2000;

259. Nakayama R, Arikawa K, Bhawal UK. The epigenetic regulation of CXCL14 plays a role in the pathobiology of oral cancers. J Cancer. 2017;

260. Kaskas NM, Moore-Medlin T, McClure GB, Ekshyyan O, Vanchiere JA, Nathan CAO. Serum biomarkers in head and neck squamous cell cancer. JAMA Otolaryngol - Head Neck Surg. 2014;

261. Chang KP, Chang YT, Wu CC, Liu YL, Chen MC, Tsang NM, et al. Multiplexed immunobead-based profiling of cytokine markers for detection of nasopharyngeal carcinoma and prognosis of patient survival. Head Neck. 2011;

262. Rentoft M, Coates PJ, Loljung L, Wilms T, Laurell G, Nylander K. Expression of CXCL10 is

associated with response to radiotherapy and overall survival in squamous cell carcinoma of the tongue. Tumor Biol. 2014;

263. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett. 2006;

264. Jablonska E, Piotrowski L, Grabowska Z. Serum Levels of IL-lβ, IL-6, TNF-α, sTNF-RI and CRP in Patients with oral cavity cancer. Pathol Oncol Res. 1997;

265. Riedel F, Zaiss I, Herzog D, Götte K, Naim R, Hörmann K. Serum levels of interleukin-6 in patients with primary head and neck squamous cell carcinoma. Anticancer Res. 2005;

266. Tartour E, Deneux L, Mosseri V, Jaulerry C, Brunin F, Point D, et al. Soluble interleukin-2 receptor serum level as a predictor of locoregional control and survival for patients with head and neck carcinoma: Results of a multivariate prospective study. Cancer. 1997;

267. Warnakulasuriya S, Johnson NW, Van Der Waal I. Nomenclature and classification of potentially malignant disorders of the oral mucosa. Journal of Oral Pathology and Medicine. 2007.

268. Punyani SR, Sathawane RS. Salivary level of interleukin-8 in oral precancer and oral squamous cell carcinoma. Clin Oral Investig. 2013;

269. K. R, G. N, R. R, P. R, A.R. K, S.N. A. Validation of the diagnostic utility of salivary interleukin 8 in the differentiation of potentially malignant oral lesions and oral squamous cell carcinoma in a region with high endemicity. Oral Surg Oral Med Oral Pathol Oral Radiol. 2014;

270. Sahingur SE, Yeudall WA. Chemokine function in periodontal disease and oral cavity cancer. Frontiers in Immunology. 2015.

271. Irani S, Barati I, Badiei M. Periodontitis Periodontitis and oral cancer - current concepts of the etiopathogenesis. Oncol Rev. 2020;

272. Okada H, Murakami S. Cytokine expression in periodontal health and disease. Critical Reviews in Oral Biology and Medicine. 1998.

273. Varki A, Kannagi R, Toole B, Stanley P. Glycosylation changes in cancer. Essentials Glycobiol 3rd Ed. 2017;

274. Shu J, Yu H, Li X, Zhang D, Liu X, Du H, et al. Salivary glycopatterns as potential biomarkers for diagnosis of gastric cancer. Oncotarget. 2017;

275. Liu X, Yu H, Qiao Y, Yang J, Shu J, Zhang J, et al. Salivary Glycopatterns as Potential Biomarkers for Screening of Early-Stage Breast Cancer. EBioMedicine. 2018;

276. Varelas X, Bouchie MP, Kukuruzinska MA. Protein N-glycosylation in oral cancer: Dysregulated cellular networks among DPAGT1, E-cadherin adhesion and canonical Wnt signaling. Glycobiology. 2014.

277. Guile GR, Harvey DJ, O'Donnell N, Powell AK, Hunter AP, Zamze S, et al. Identification of highly fucosylated N-linked oligosaccharides from the human parotid gland. Eur J Biochem. 1998;

278. Everest-Dass A V., Jin D, Thaysen-Andersen M, Nevalainen H, Kolarich D, Packer NH. Comparative structural analysis of the glycosylation of salivary and buccal cell proteins: Innate protection against infection by Candida albicans. Glycobiology. 2012;

279. Sinevici N, Mittermayer S, Davey GP, Bones J, O'Sullivan J. Salivary N-glycosylation as a

biomarker of Oral Cancer: A pilot study. Glycobiology. 2019;

280. Blanas A, Sahasrabudhe NM, Rodríguez E, van Kooyk Y, van Vliet SJ. Fucosylated antigens in cancer: An alliance toward tumor progression, metastasis, and resistance to chemotherapy. Frontiers in Oncology. 2018.

281. Lin WL, Lin YS, Shi GY, Chang CF, Wu HL. Lewisy promotes migration of oral cancer cells by glycosylation of epidermal growth factor receptore0120162. PLoS One. 2015;

282. Miyoshi E, Moriwaki K, Nakagawa T. Biological function of fucosylation in cancer biology. Journal of Biochemistry. 2008.

283. Shah M, Telang S, Raval G, Shah P, Patel PS. Serum fucosylation changes in oral cancer and oral precancerous conditions: α-L-fucosidase as a marker. Cancer. 2008;

284. Kumar S, Suhag A, Kolay S, Kumar P, Narwal A, Srinivas K, et al. Serum fucose level in oral cancer, leukoplakia, and oral sub mucous fibrosis: A biochemical study. J Fam Med Prim Care. 2019;

285. Y. M, A. H, S. B, V. P, M. Z, L.E. D, et al. Quantitative serum glycomics of esophageal adenocarcinoma and other esophageal disease onsets. Journal of Proteome Research. 2009.

286. Chen YT, Chong YM, Cheng CW, Ho CL, Tsai HW, Kasten FH, et al. Identification of novel tumor markers for oral squamous cell carcinoma using glycoproteomic analysis. Clin Chim Acta. 2013;

287. Lu J, Gu J. Significance of β-galactoside α2,6 sialyltranferase 1 in cancers. Molecules. 2015.

288. Achalli S, Madi M, Babu SG, Shetty SR, Kumari S, Bhat S. Sialic acid as a biomarker of oral potentially malignant disorders and oral cancer. Indian J Dent Res. 2017;

289. Liu H, Sadygov RG, Yates JR. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem. 2004;

290. Hu S, Loo JA, Wong DT. Human saliva proteome analysis and disease biomarker discovery. Expert Review of Proteomics. 2007.

291. Krapfenbauer K, Drucker E, Thurnher D. Identification of tumor-related proteins as potential screening markers by proteome analysis-protein profiles of human saliva as a predictive and prognostic tool. EPMA J. 2014;

292. Khan RS, Khurshid Z, Akhbar S, Moin SF. Advances of salivary proteomics in oral squamous cell carcinoma (OSCC) detection: An update. Proteomes. 2016.

293. Shintani S, Hamakawa H, Ueyama Y, Hatori M, Toyoshima T. Identification of a truncated cystatin SA-I as a saliva biomarker for oral squamous cell carcinoma using the SELDI ProteinChip platform. Int J Oral Maxillofac Surg. 2010;

294. Wu CC, Chu HW, Hsu CW, Chang KP, Liu HP. Saliva proteome profiling reveals potential salivary biomarkers for the detection of oral cavity squamous cell carcinoma. Proteomics. 2015;

295. Csősz É, Márkus B, Darula Z, Medzihradszky KF, Nemes J, Szabó E, et al. Salivary proteome profiling of oral squamous cell carcinoma in a Hungarian population. FEBS Open Bio. 2018;

296. Magnussen S, Rikardsen OG, Hadler-Olsen E, Uhlin-Hansen L, Steigen SE, Svineng G. Urokinase Plasminogen Activator Receptor (uPAR) and Plasminogen Activator Inhibitor-1 (PAI-1) are potential predictive biomarkers in early-stage Oral Squamous Cell Carcinomas

(OSCC). PLoS One. 2014;

297. de Carvalho Fraga CA, Farias LC, Jones KM, Batista de Paula AM, Guimaraes ALS. Angiotensin-Converting Enzymes (ACE and ACE2) as Potential Targets for Malignant Epithelial Neoplasia: Review and Bioinformatics Analyses Focused in Oral Squamous Cell Carcinoma. Protein Pept Lett. 2017;

298. Tung CL, Lin ST, Chou HC, Chen YW, Lin HC, Tung CL, et al. Proteomics-based identification of plasma biomarkers in oral squamous cell carcinoma. J Pharm Biomed Anal. 2013;

299. Papa F, Siciliano RA, Inchingolo F, Mazzeo MF, Scacco S, Lippolis R. Proteomics pattern associated with gingival oral squamous cell carcinoma and epulis: A case analysis. Oral Science International. 2018.

300. Jou YJ, Hua CH, Lin C Der, Lai CH, Huang SH, Tsai MH, et al. S100A8 as potential salivary biomarker of oral squamous cell carcinoma using nanoLC-MS/MS. Clin Chim Acta. 2014;

301. Ralhan R, DeSouza L V., Matta A, Tripathi SC, Ghanny S, Datta Gupta S, et al. Discovery and verification of head-and-neck cancer biomarkers by differential protein expression analysis using iTRAQ labeling, multidimensional liquid chromatography, and tandem mass spectrometry. Mol Cell Proteomics. 2008;

302. Kesting MR, Sudhoff H, Hasler RJ, Nieberler M, Pautke C, Wolff KD, et al. Psoriasin (S100A7) up-regulation in oral squamous cell carcinoma and its relation to clinicopathologic features. Oral Oncol. 2009;

303. Tabatabaeifar S, Thomassen M, Larsen MJ, Larsen SR, Kruse TA, Sørensen JA. The subclonal structure and genomic evolution of oral squamous cell carcinoma were revealed by ultra-deep sequencing. Oncotarget. 2017;

304. Scharenberg C, Eckardt A, Tiede C, Kreipe H, Hussein K. Expression of Caspase 14 and Filaggrin in Oral Squamous Carcinoma. Head Neck Pathol. 2013;

305. Liu C, Tang J, Li C, Pu G, Yang D, Chang X. PADI4 stimulates esophageal squamous cell carcinoma tumor growth and up-regulates CA9 expression. Mol Carcinog. 2019;

306. Lee CW, Chang KP, Chen YY, Liang Y, Hsueh C, Yu JS, et al. Overexpressed tryptophanyl-tRNA synthetase, an angiostatic protein, enhances oral cancer cell invasiveness. Oncotarget. 2015;

307. Mazzoccoli G, Castellana S, Carella M, Palumbo O, Tiberio C, Fusilli C, et al. A primary tumor gene expression signature identifies a crucial role played by tumor stroma myofibroblasts in lymph node involvement in oral squamous cell carcinoma. Oncotarget. 2017;

308. Syed N, Chavan S, Sahasrabuddhe NA, Renuse S, Sathe G, Nanjappa V, et al. Silencing of high-mobility group box 2 (HMGB2) modulates cisplatin and 5-fluorouracil sensitivity in head and neck squamous cell carcinoma. Proteomics. 2015;

309. Jeanes A, Gottardi CJ, Yap AS. Cadherins and cancer: How does cadherin dysfunction promote tumor progression? Oncogene. 2008.

310. Mandal M, Myers JN, Lippman SM, Johnson FM, Williams MD, Rayala S, et al. Epithelial to mesenchymal transition in head and neck squamous carcinoma: Association of Src activation with E-cadherin down-regulation, vimentin expression, and aggressive tumor features. Cancer. 2008;

311. Liu LK, Jiang XY, Zhou XX, Wang DM, Song XL, Jiang HB. Upregulation of vimentin and aberrant expression of E-cadherin/B- catenin complex in oral squamous cell carcinomas:

Correlation with the clinicopathological features and patient outcome. Mod Pathol. 2010;

312. Adami GR, O'Callaghan TN, Kolokythas A, Cabay RJ, Zhou Y, Schwartz JL. A loss of profilin-1 in late-stage oral squamous cell carcinoma. J Oral Pathol Med. 2017;

313. Yu Q, Xu L, Chen L, Sun B, Yang Z, Lu K, et al. Vinculin expression in non-small cell lung cancer. J Int Med Res. 2019;

314. Pedro NF, Biselli JM, Maniglia JV, Santi-Neto D de, Pavarino ÉC, Goloni-Bertollo EM, et al. Candidate biomarkers for oral squamous cell carcinoma: Differential expression of oxidative stress-related genes. Asian Pacific J Cancer Prev. 2018;

315. Miyashita H, Uchida T, Mori S, Echigo S, Motegi K. Expression status of Pin1 and cyclins in oral squamous cell carcinoma: Pin1 correlates with Cyclin D1 mRNA expression and clinical significance of cyclins. Oncol Rep. 2003;

316. Meram AT, Chen J, Patel S, Kim DD, Shirley B, Covello P, et al. Hydrogen sulfide is increased in oral squamous cell carcinoma compared to adjacent benign oral mucosae. Anticancer Res. 2018;

317. Sasahira T, Kurihara M, Nishiguchi Y, Nakashima C, Kirita T, Kuniyasu H. Pancreatic adenocarcinoma up-regulated factor has oncogenic functions in oral squamous cell carcinoma. Histopathology. 2017;

318. Sahni A, Rehani S, Kardam P, Sethi S, Kumari R, Mathias Y. Analysis of stromal mucin in oral epithelial dysplasia & oral squamous cell carcinoma- A histochemical study. J Oral Biol Craniofacial Res. 2019;

319. Kinoshita T, Nohata N, Watanabe-Takano H, Yoshino H, Hidaka H, Fujimura L, et al. Actin-related protein 2/3 complex subunit 5 (ARPC5) contributes to cell migration and invasion and is directly regulated by tumor-suppressive microRNA-133a in head and neck squamous cell carcinoma. Int J Oncol. 2012;

320. Tekkeşin MS, Mutlu S, Aksakalli N, Olgaç V. Expression of heat shock proteins 27, 60 and 70 in oral carcinogenesis: An immunohistochemical study. Turk Onkol Derg. 2011;

321. Kawasaki G, Kato Y, Mizuno A. Cathepsin expression in oral squamous cell carcinoma: Relationship with clinicopathologic factors. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2002;

322. Zimmermann BG, Wong DT. Salivary mRNA targets for cancer diagnostics. Oral Oncology. 2008.

323. Manchil PRD, Joy ET, Kiran MS, Sherubin JE, Khan MF, Aravind BS. Correlation of serum levo-fucose levels as a biomarker with tumor node metastasis staging in oral cancer patients. J Pharm Bioallied Sci. 2016;

324. Mei CM, Shen S. The roles of cell adhesion molecules in tumor suppression and cell migration: A new paradox. Cell Adhesion and Migration. 2009.

325. Gao Y, Wang Z, Hao Q, Li W, Xu Y, Zhang J, et al. Loss of ERα induces amoeboid-like migration of breast cancer cells by downregulating vinculin. Nat Commun. 2017;

326. Li T, Guo H, Song Y, Zhao X, Shi Y, Lu Y, et al. Loss of vinculin and membrane-bound β-catenin promotes metastasis and predicts poor prognosis in colorectal cancer. Mol Cancer. 2014;

327. Yu JS, Chen YT, Chiang WF, Hsiao YC, Chu LJ, Seei LC, et al. Saliva protein biomarkers To detect oral squamous cell carcinoma in a high-risk population in Taiwan. Proc Natl Acad Sci

U S A. 2016;

328. Pettus JR, Johnson JJ, Shi Z, Davis JW, Koblinski J, Ghosh S, et al. Multiple kallikrein (KLK 5, 7, 8, and 10) expression in squamous cell carcinoma of the oral cavity. Histol Histopathol. 2009;

329. Yokoe H, Nomura H, Yamano Y, Fushimi K, Sakamoto Y, Ogawara K, et al. Alteration of extracellular superoxide dismutase expression is associated with an aggressive phenotype of oral squamous-cell carcinoma. Exp Ther Med. 2010;

330. Narayana N, Gist J, Smith T, Tylka D, Trogdon G, Wahl JK. Desmosomal component expression in normal, dysplastic, and oral squamous cell carcinoma. Dermatol Res Pract. 2010;

331. Xu W, Zou H, Wei Z, Song C, Tang C, Yin X, et al. Rh type C-glycoprotein functions as a novel tumor suppressor gene by inhibiting tumorigenicity and metastasis in head and neck squamous cell carcinoma. Aging (Albany NY). 2019;

332. Fillies T, Jogschies M, Kleinheinz J, Brandt B, Joos U, Buerger H. Cytokeratin alteration in oral leukoplakia and oral squamous cell carcinoma. Oncol Rep. 2007;

333. Villaret DB, Wang T, Dillon D, Xu J, Sivam D, Cheever MA, et al. Identification of genes overexpressed in head and neck squamous cell carcinoma using a combination of complementary DNA subtraction and microarray analysis. Laryngoscope. 2000;

334. Sakamoto K, Aragaki T, Morita K ichi, Kawachi H, Kayamori K, Nakanishi S, et al. Down-regulation of keratin 4 and keratin 13 expression in oral squamous cell carcinoma and epithelial dysplasia: A clue for histopathogenesis. Histopathology. 2011;

335. Feller L, Altini M, Lemmer J. Inflammation in the context of oral cancer. Oral Oncology. 2013.

336. Feng Y, Li Q, Chen J, Yi P, Xu X, Fan Y, et al. Salivary protease spectrum biomarkers of oral cancer. Int J Oral Sci. 2019;

VIII. *ABBREVIATION INDEX*

ABC - Ammonium Bicarbonate

AC - Actinic Cheilitis

AFP - α-Fetoprotein

AREG - Amphiregulin

BCA - Bicinchoninic acid assay

Bcl-2 - B-Cell Lymphoma-2

BSCC - Basaloid SCC

BTC - Betacellulin

CA125 - Cancer Antigen 125

CA15-3 - Cancer Antigen 15-3

CCD1 - Cyclin D1

CDK- Cyclin-Dependent kinase

CDKN2a - Cyclin-Dependent Kinase inhibitor 2a

CIS - Carcinoma *In Situ*

c-MET - Mesenchymal Epithelial Transition factor

CSC - Cancer Stem Cell

CT - Computer Tomography

DC - Dendritic cells

DNA - Deoxyribonucleic Acid

DTT- Dithiothreitol

EBV - Epstein-Barr Virus

EGF - Epidermal Growth Factor

EGFR - Epidermal Growth Factor Receptor

ELISA - Enzyme-Linked Immunosorbent Assay

EMT - Epithelial-Mesenchymal Transition

EPGN - Epigen

EREG - Epiregulin

FDA - Food and Drug Agency

FFPET - Formalin-fixed paraffin-embedded tissue

FGF1/2 - Fibroblast Growth Factor 1/2 gamma

GLANT2 - acetylgalactosaminyltransferase 2

HBEGF - Heparin-Binding EGF-like growth factor

HCC-1 -Hemofiltrate CC chemokine 1

HDL - High-Density Lipoprotein

HE4 - Human Epididymis protein 4

HER2 - Human Epidermal growth factor Receptor 2

HGF - Hepatocyte Growth Factor

HL - Homogeneous Leukoplakia

HNC - Head and Neck Cancer

HPV - Human Papilloma Virus

hTERC - human Telomerase RNA gene

IAA - Iodoacetamide

IAP - Inhibitor of Apoptosis

IARC - International Agency for Research on Cancer

IFNs - Interferons

IFN-γ - Interferon **IHC** - Immunohistochemistry

IgA – Immunoglobulin A

ILs - Interleukins

IP - interferon gamma-induced protein

Jak - Janus kinase

kD - kilodalton

LC - Liquid Chromatography

LP- Lichen Planus

LT-α - Lymphotoxin-alpha

MAPK - Mitogen-activated Protein Kinase

MCP- Monocyte chemoattractant protein

MF - Macrophages

MHC - Major Histocompatibility Complex

MRI - Magnetic Resonance Imaging

MS - Mass Spectrometry

mTOR - mammalian Target of Rapamycin

MVD - Microvessel Density

MVs - Microvesicles

MWCO - Molecular Weight Cut-Off

NF- κB - Nuclear Factor kappa-light-chain-enhancer of activated B cells

NSCs - Normal Stem Cells

OL - Oral Leukoplakia

OLP - Oral Lichen Planus

OPMDs - Oral Potentially Malignant Disorders

OPSCC - Oropharyngeal Squamous Cell Carcinoma

OSCC - Oral Squamous Cell Carcinoma

OSMF - Oral Submucous Fibrosis

PAGE Polyacrylamide Gel Electrophoresis

PD-1 - Programmed cell death protein 1

PD-L1 - Programmed cell death protein 1 ligand

PET - Positron Emission Tomography

PF-4 - Platelet factor 4

PI3K - PhosphatIdylinositol-3- kinase

PKC - Protein Kinase C

PNGaseF - Peptide N-glycosidase

pRb - Retinoblastoma protein

PSA - Porstate-Specific Antigen

PTMs - Posttranslational modifications

PVL - Proliferative Verrucous Leukoplakia

RNA - Ribonucleic Acid

SCC - Squamous Cell Carcinoma

SF - Scatter Factor

SOX2 - Sex determining region Y-box 2

STAT - Signal Transducer and Activator of Transcription

STK - Smokeless Tobacco Keratosis

TACs - Transit Amplifying Cells

TAMs - Tumour-Associated Macrophages

TC - Total Cholesterol

Tg - Thyroglobulin

TGFα- Transforming growth Factor-alpha

TGF-β - Transforming Growth Factor beta

TNF-α - tumor necrosis factors alpha

TNM - Tumour- Lymph Node-Metastasis

UHPLC - Ultra-High-Performance Liquid Chromatography

VC - Verrucous Carcinoma

VLDL - Very low-density lipoprotein

WHO - World Health Organization

IX. *SUPPLEMENTARY FIGURES*

Figure 28. Dot plots showing log2 concentration (pg/mL) of salivary (**A**) IL-1α, (**B**) IL-6, (**C**) IL-8, (**D**) IP-10, (**E**) MCP-1, (**F**)TNF-α, (**G**) HCC-1 and (**H**) PF-4 s in HL and PVL patients. Values represent mean ±SEM of n=33 (HL) and n=33 (PVL), where n is an average of two technical replicates.

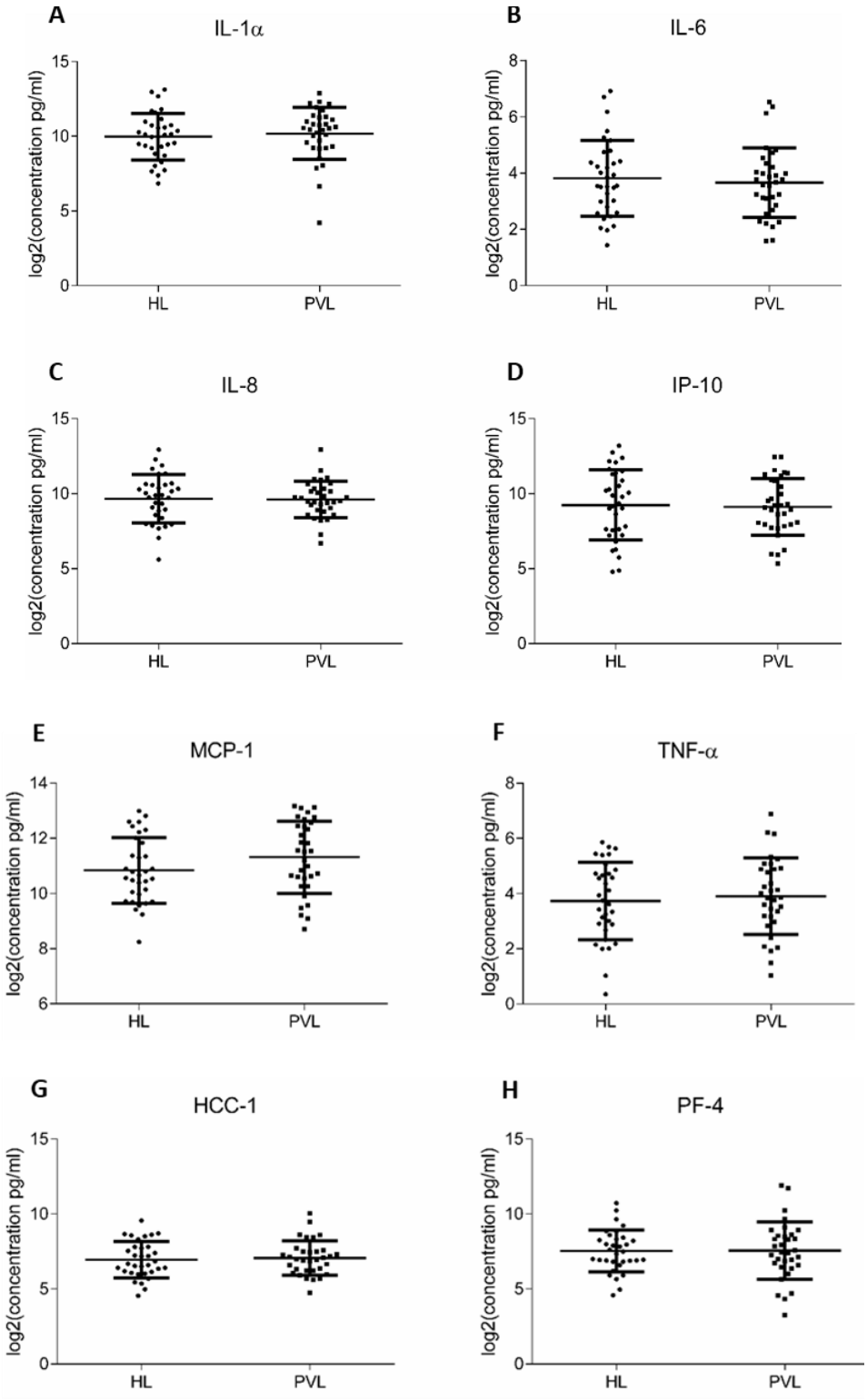
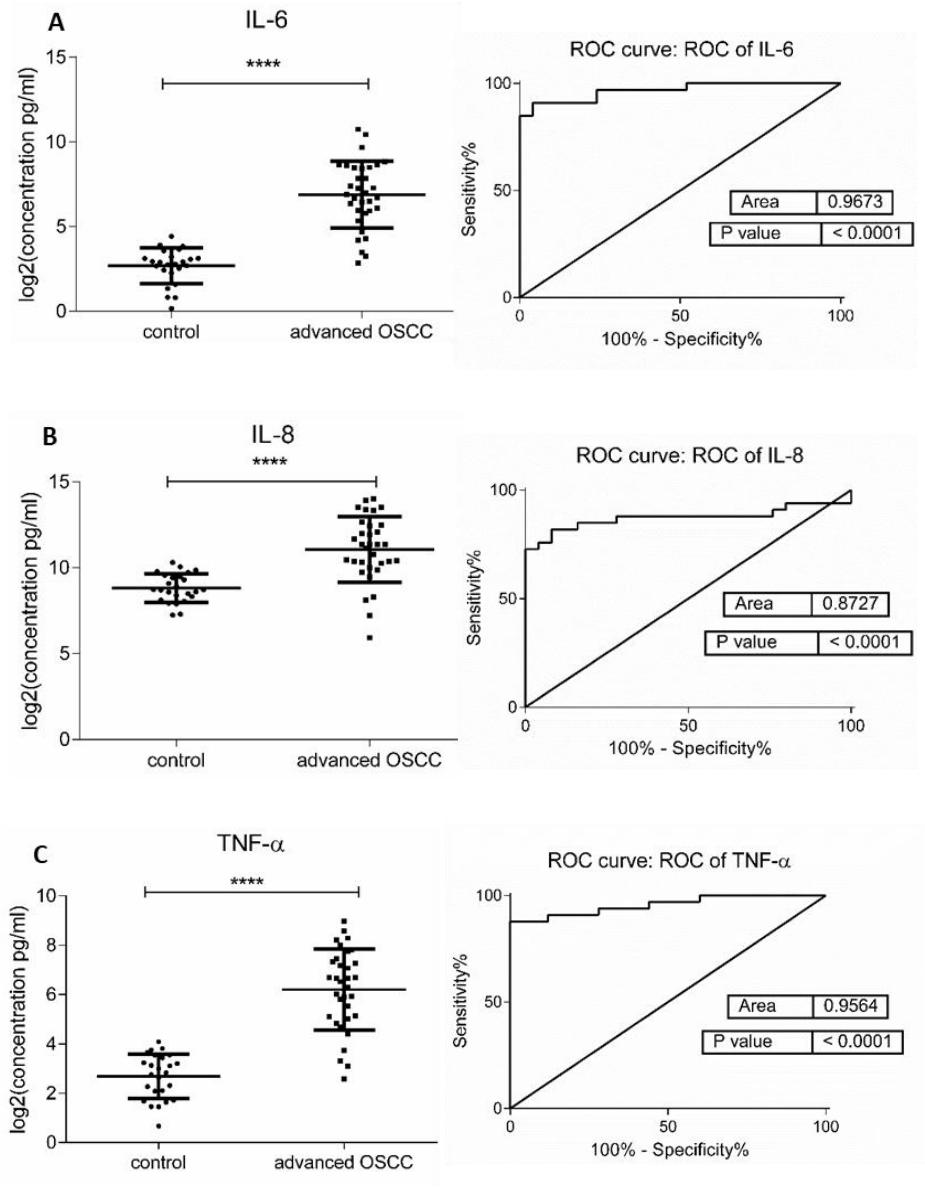
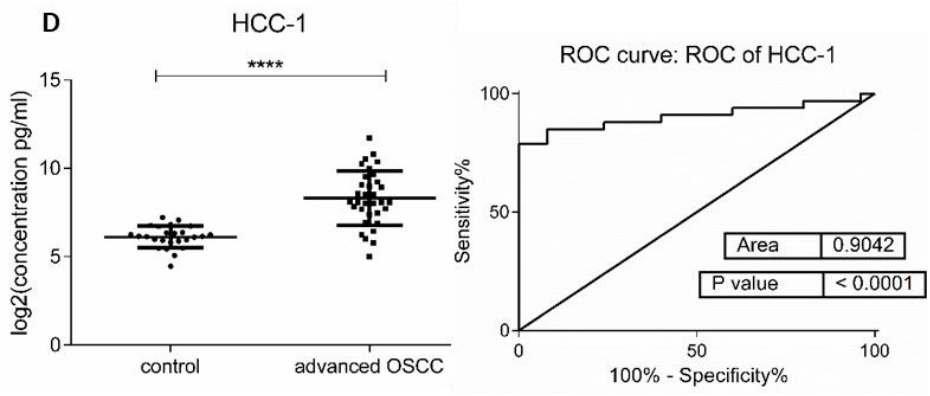
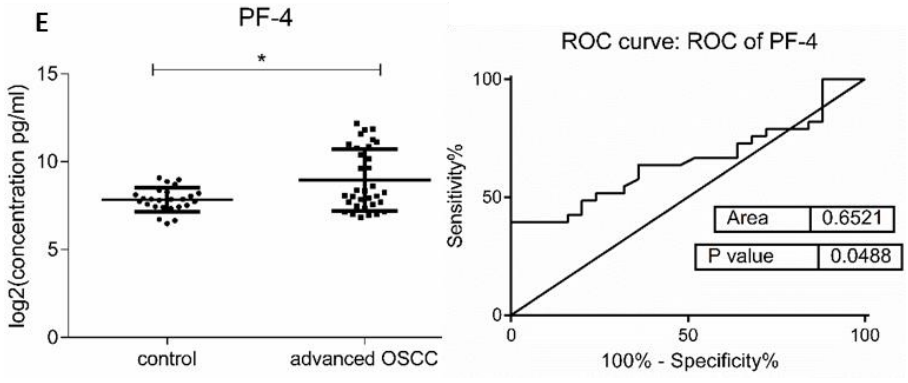


Figure 30. Cytokines expressions in patients at advanced OSCC stages and control individuals. Dot plot (left) of salivary cytokine log2 levels (pg/mL) and ROC curves (right) of (**A**) IL-6, (**B**) IL-8, (**C**) TNF-α, (**D**) HCC-1 and (**E**) PF-4 (p = 0.0001, 0.0001, 0.0001, 0.0001 and 0.001, respectively). Dot plots of (**F**) IL-α, (**G**) IP-10 and (**H**) MCP-1. Values represent mean ±SEM of n=25 (control) and n=33 (advanced OSCC), where n is an average of two technical replicates.




D
HCC-1
****
log2(concentration pg/ml)
15
10
5
0
control
advanced OSCC
ROC curve: ROC of HCC-1
Sensitivity%
100
50
0
0
50
100
100% - Specificity%
Area
0.9042
P value
< 0.0001


E
PF-4
*
log2(concentration pg/ml)
15
10
5
0
control
advanced OSCC
ROC curve: ROC of PF-4
Sensitivity%
100
50
0
0
50
100
100% - Specificity%
Area
0.6521
P value
0.0488

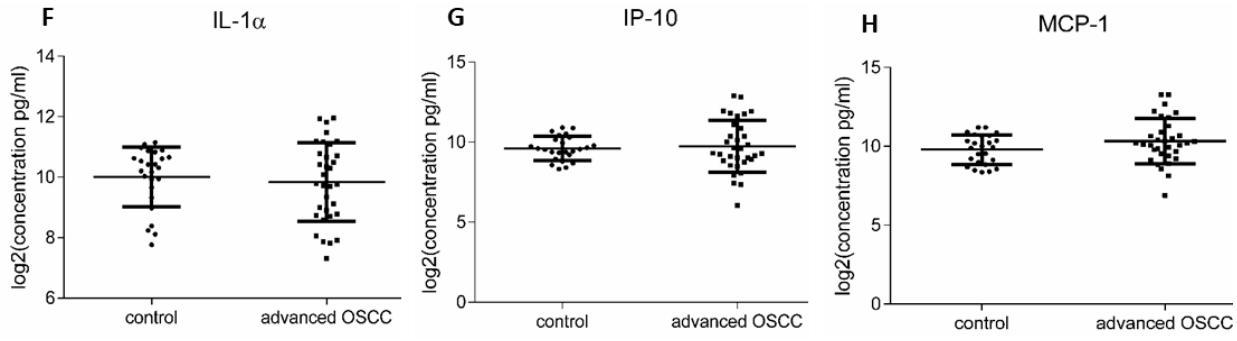
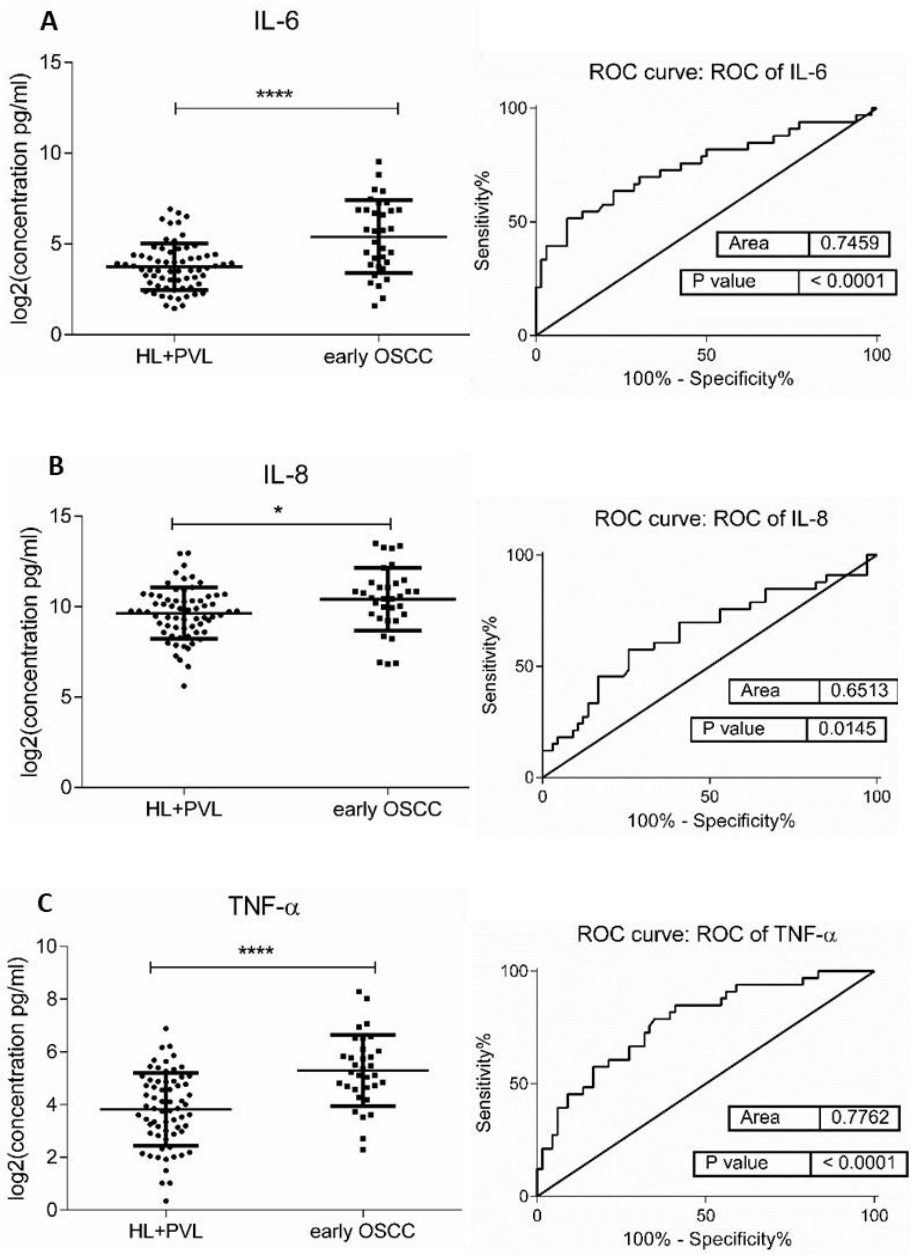
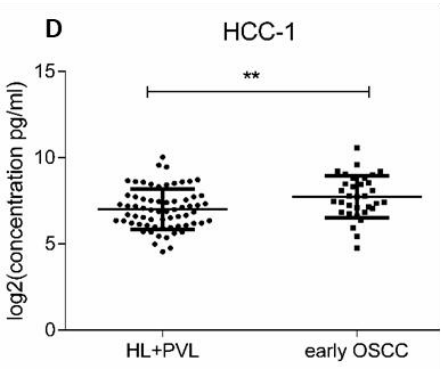
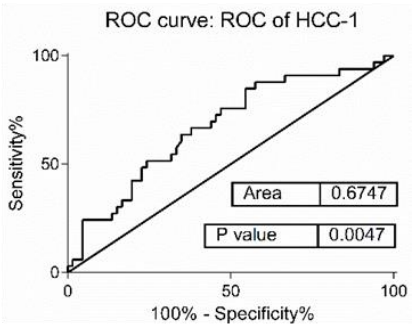
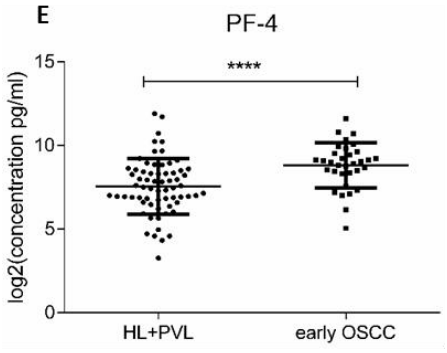

F
IL-1α
log2(concentration pg/ml)
14
12
10
8
6
control
advanced OSCC
G
IP-10
log2(concentration pg/ml)
15
10
5
0
control
advanced OSCC
H
MCP-1
log2(concentration pg/ml)
15
10
5
0
control
advanced OSCC

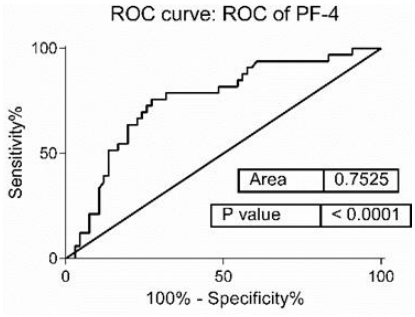
Figure 32. Comparison of salivary inflammatory factors in patients with oral potentially malignant lesions (HL&PVL) and early OSCC stages. Dot plot (left) and ROC curve (right) showing protein log2 levels (pg/mL) of (**A**) IL-6, (**B**) IL-8, (**C**) TNF-α, (**D**) HCC-1 and (**E**) PF-4 (p = 0.0001, 0.05, 0.0001, 0.01 and 0.0001, respectively). Dot plots of (**F**) IL-1α, (**G**) IP-10 and (**H**) MCP-1. Values represent mean ±SEM of n=66 (HL&PVL) and n=33 (early OSCC), where n is an average of two technical replicates.

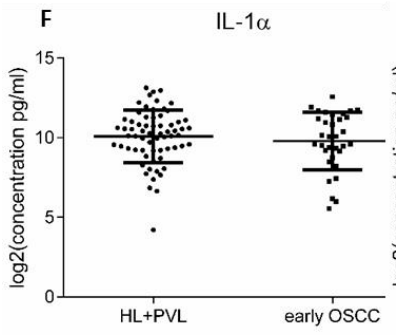


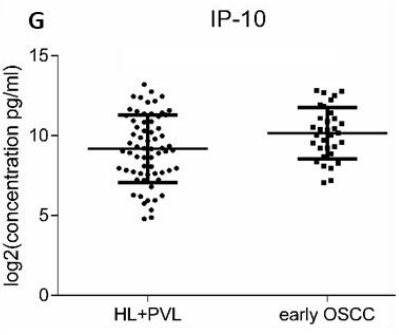

D
HCC-1
**
log2(concentration pg/ml)
15
10
5
0
HL+PVL
early OSCC


ROC curve: ROC of HCC-1
Sensitivity%
100
50
0
Area
0.6747
P value
0.0047
0
50
100
100% - Specificity%


E
PF-4
****
log2(concentration pg/ml)
15
10
5
0
HL+PVL
early OSCC


ROC curve: ROC of PF-4
Sensitivity%
100
50
0
Area
0.7525
P value
< 0.0001
0
50
100
100% - Specificity%


F
IL-1α
log2(concentration pg/ml)
15
10
5
0
HL+PVL
early OSCC


G
IP-10
log2(concentration pg/ml)
15
10
5
0
HL+PVL
early OSCC

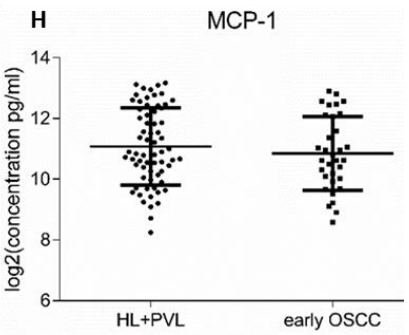
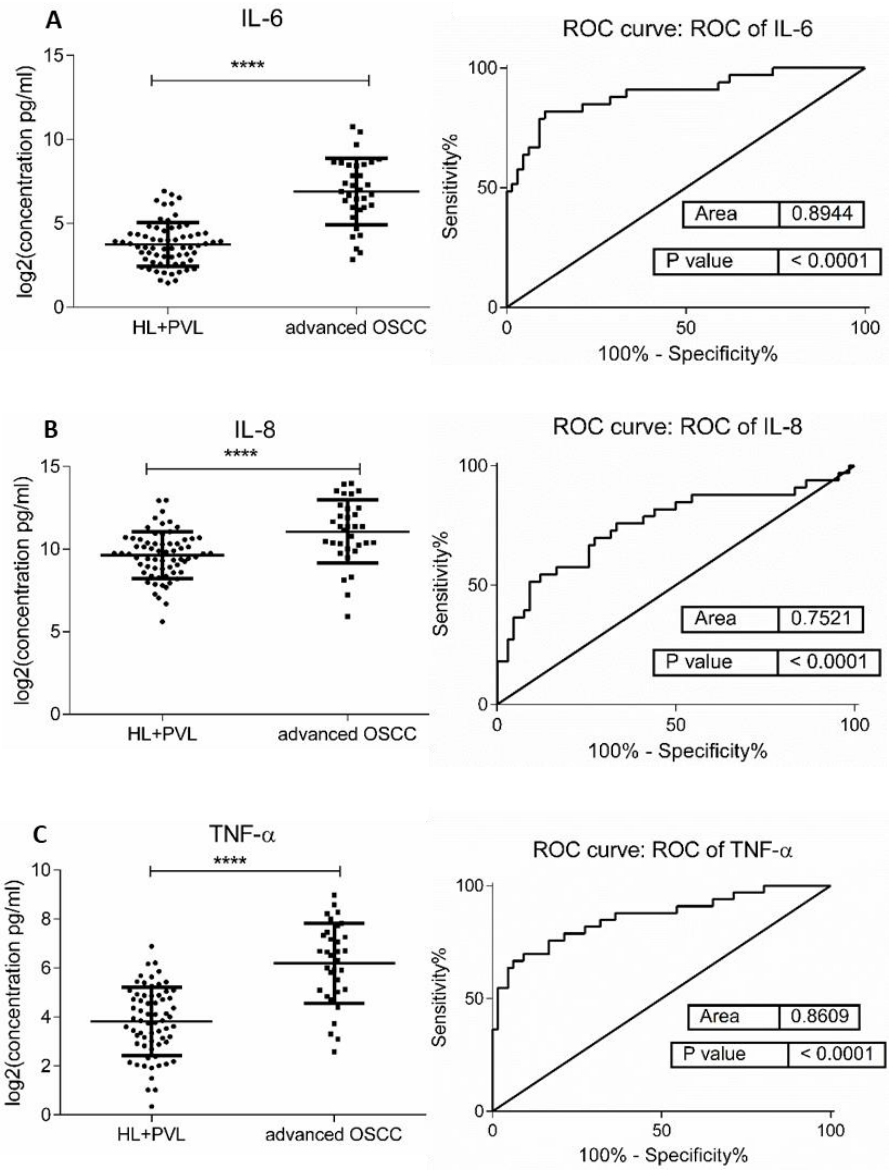
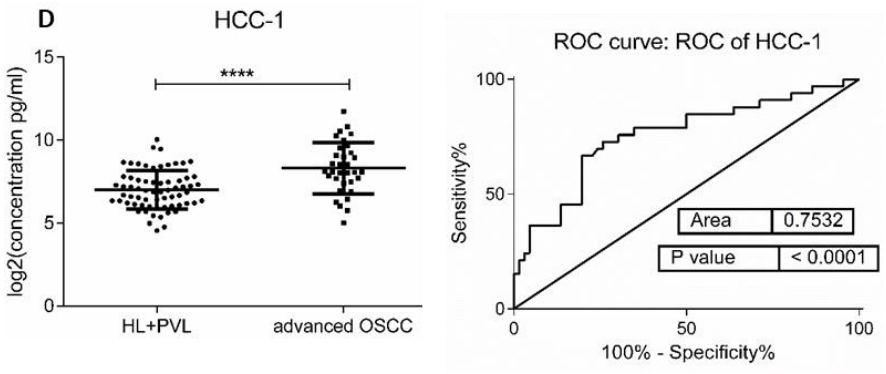
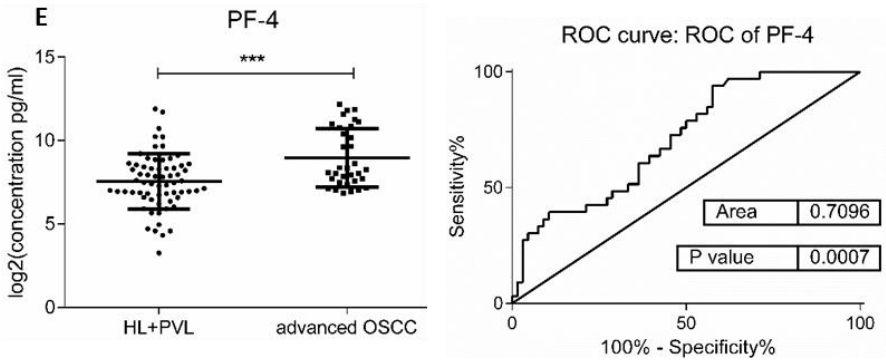

H
MCP-1
log2(concentration pg/ml)
14
12
10
8
6
HL+PVL
early OSCC

Figure 33. Inflammatory factors expression in the saliva of patients with oral potentially malignant lesions (HL&PVL) and at advanced OSCC stages. Dot plot (left) and ROC curve (right) showing cytokine log2 levels (pg/mL) of (**A**) IL-6, (**B**) IL-8, (**C**) TNF-α, (**D**) HCC-1 and (**E**) PF-4 (p = 0.0001, 0.0001, 0.0001, 0.0001 and 0.001, respectively). Dot plots of (**F**) IL-1α, (**G**) IP-10, and (**H**) MCP-1. Values represent mean ±SEM of n=66 (HL&PVL) and n=33 (advanced OSCC), where n is an average of two technical replicates.




D
HCC-1
****
log2(concentration pg/ml)
HL+PVL
advanced OSCC
ROC curve: ROC of HCC-1
Sensitivity%
100% - Specificity%
Area
0.7532
P value
< 0.0001


E
PF-4
***
log2(concentration pg/ml)
HL+PVL
advanced OSCC
ROC curve: ROC of PF-4
Sensitivity%
100% - Specificity%
Area
0.7096
P value
0.0007

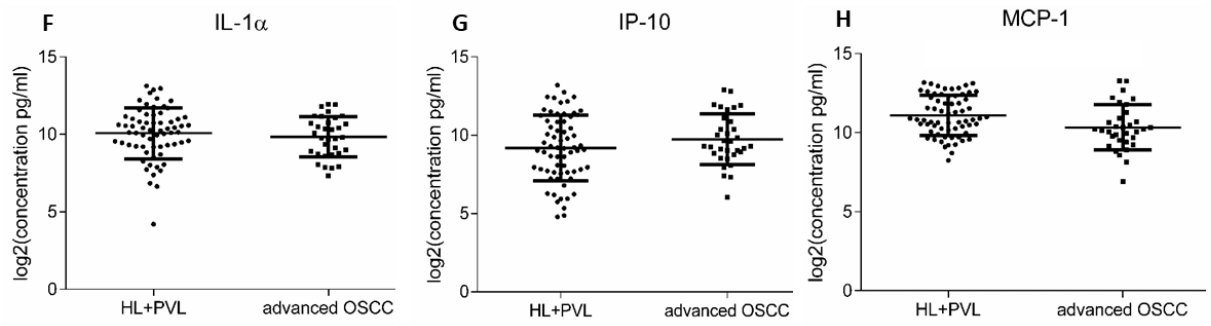

F
IL-1α
G
IP-10
H
MCP-1
log2(concentration pg/ml)
HL+PVL
advanced OSCC

Figure 49. Receiver operating characteristic (ROC) curves and corresponding area under the curve (AUC) statistics for considerably **(A)** upregulated and **(B)** downregulated salivary proteins in PVL compared to the control group. The true positive rate (sensitivity) is plotted as a function of the false positive rate (1-specificity). The area under the ROC curve is the measure of how well the model distinguishes between control and PVL.

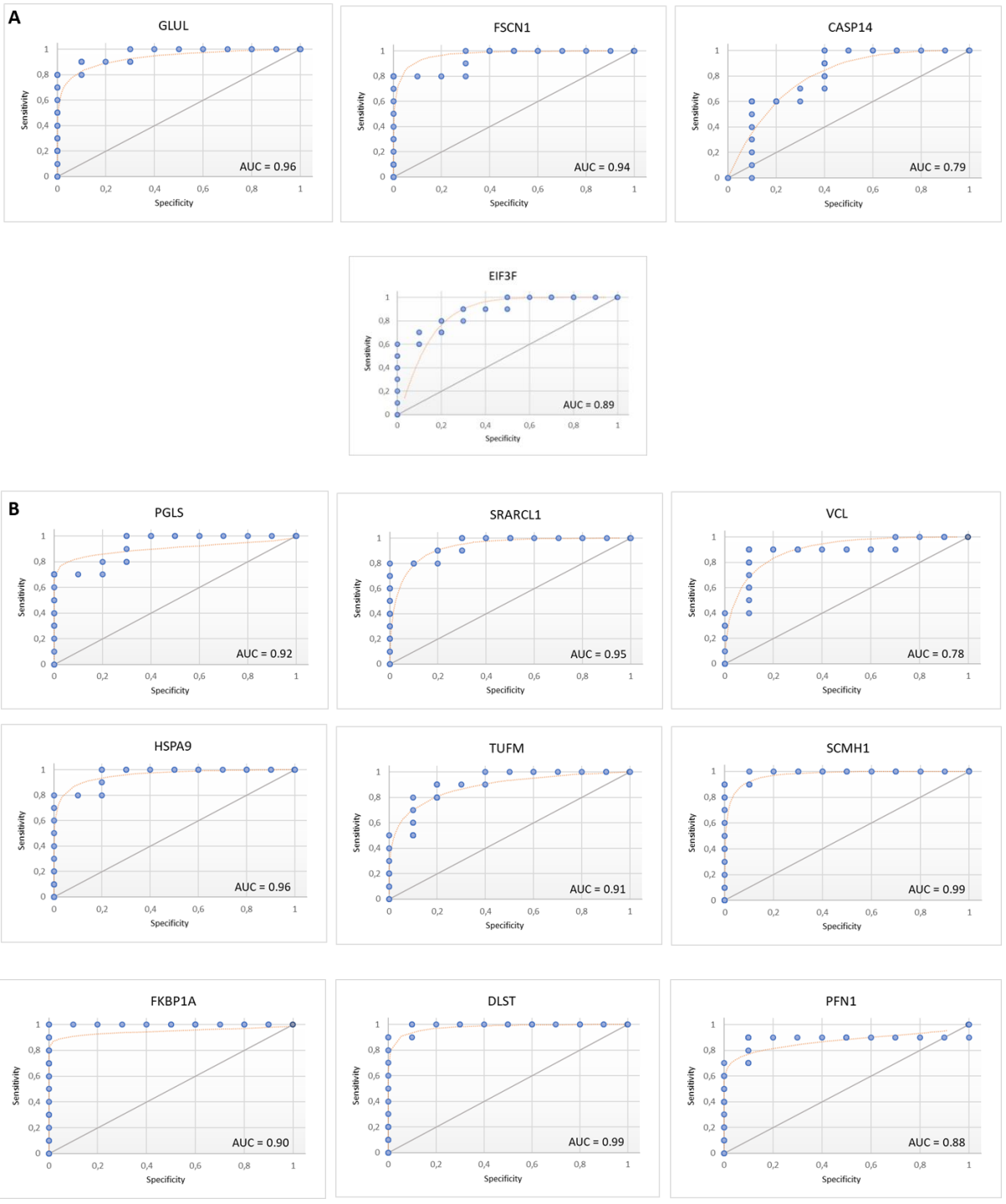
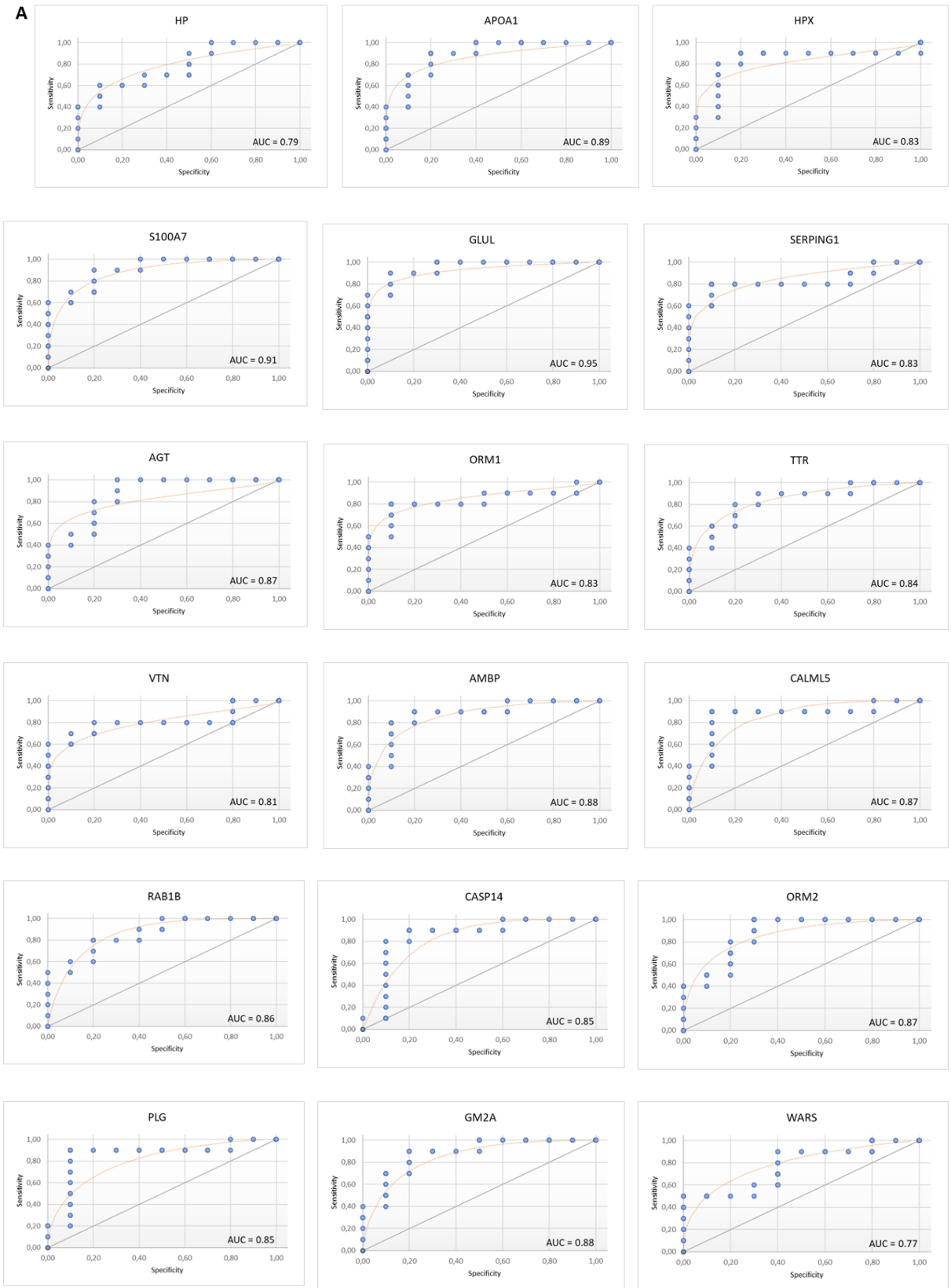


Figure 52. Receiver ROC curves and corresponding AUC statistics for salivary proteins with notably (**A**) upregulated and (**B**) down-regulated expression in early OSCC compared to control subjects.



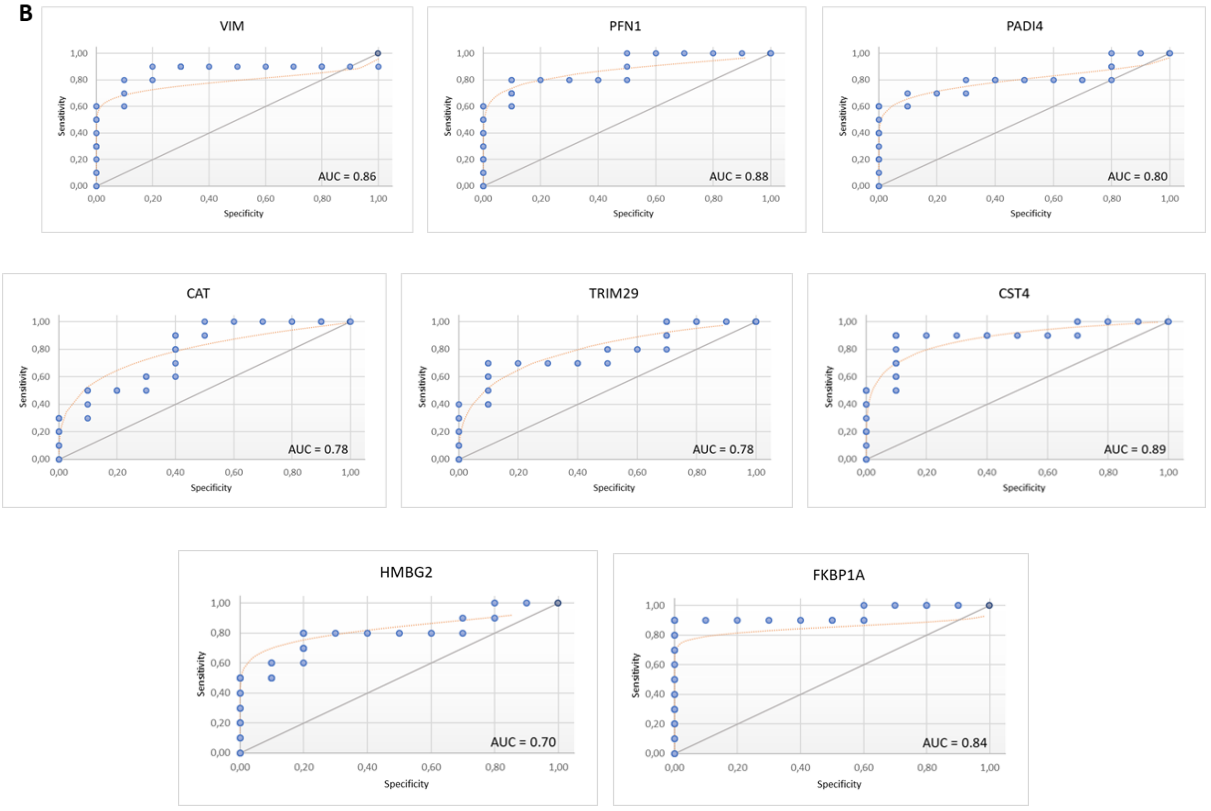
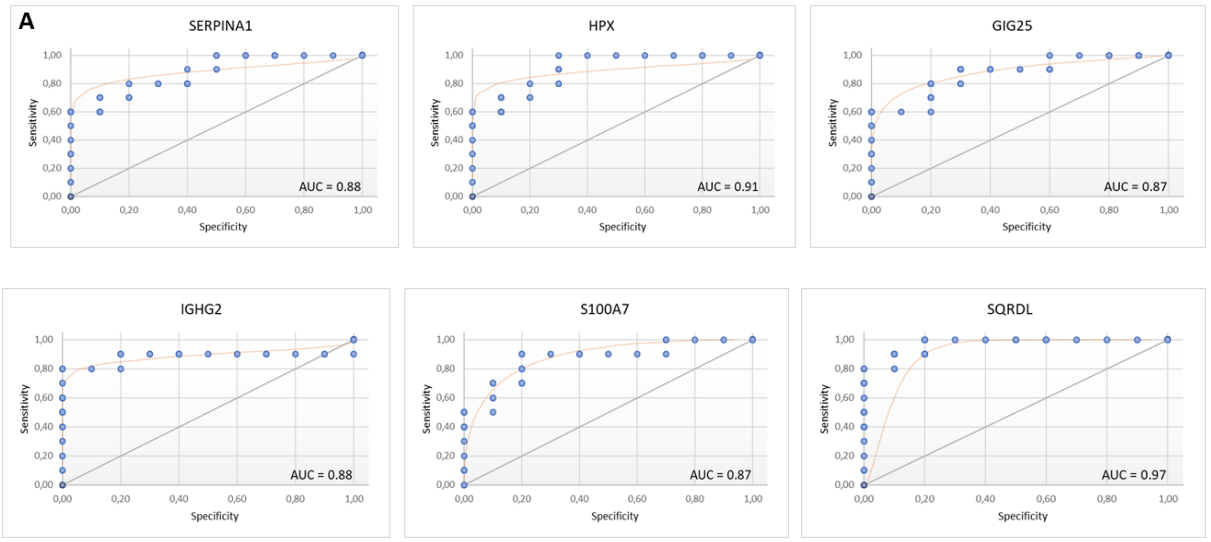
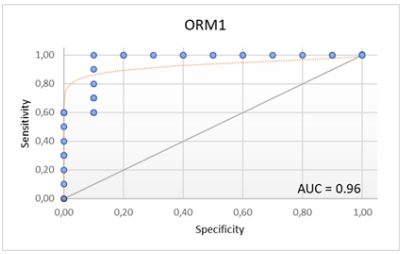
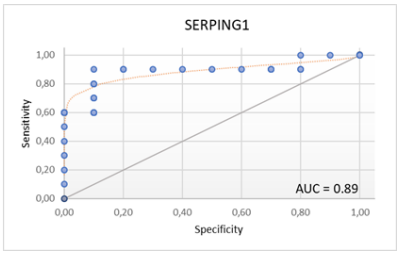
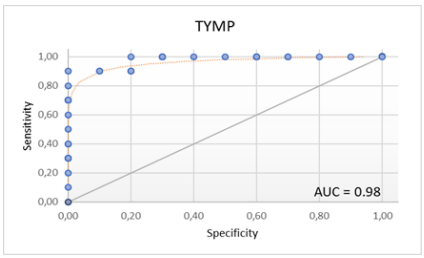


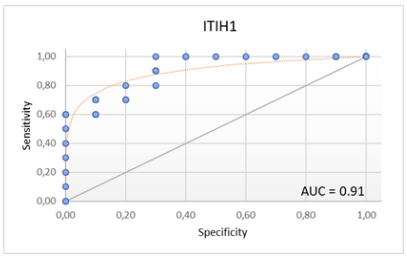
Figure 55. ROC analysis of proteins with significantly (**A**) increased and (**B**) decreased relative abundance in the saliva of patients at advanced OSCC stages compared to control individuals.



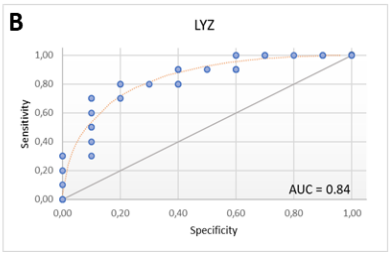

ORM1
Sensitivity
Specificity
AUC = 0.96

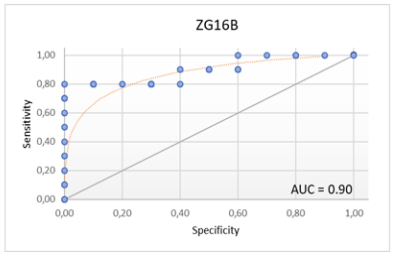

SERPING1
Sensitivity
Specificity
AUC = 0.89

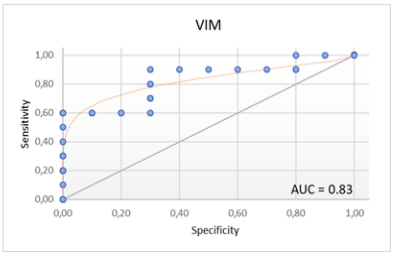

TYMP
Sensitivity
Specificity
AUC = 0.98

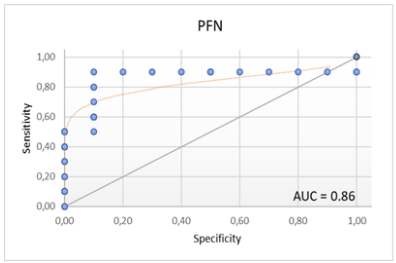

ITIH1
Sensitivity
Specificity
AUC = 0.91

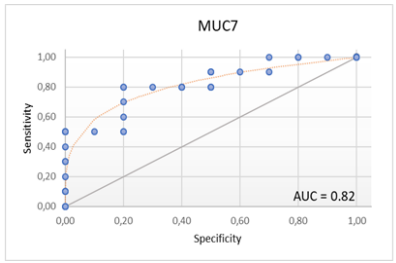
B

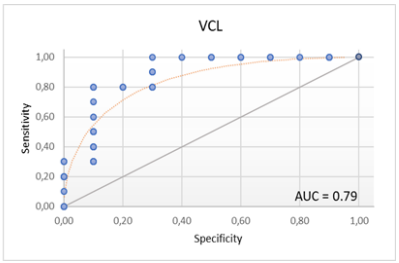

LYZ
Sensitivity
Specificity
AUC = 0.84

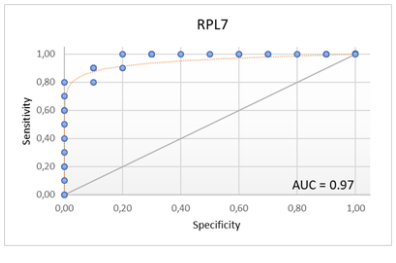

ZG16B
Sensitivity
Specificity
AUC = 0.90

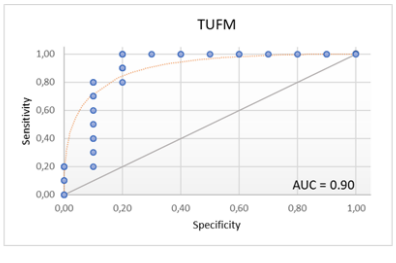

VIM
Sensitivity
Specificity
AUC = 0.83

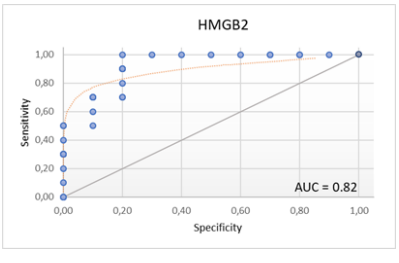

PFN
Sensitivity
Specificity
AUC = 0.86

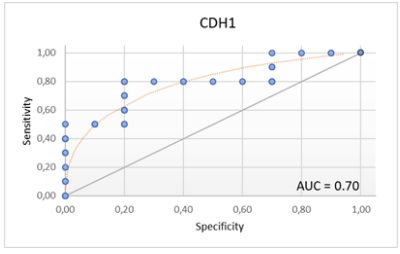

MUC7
Sensitivity
Specificity
AUC = 0.82

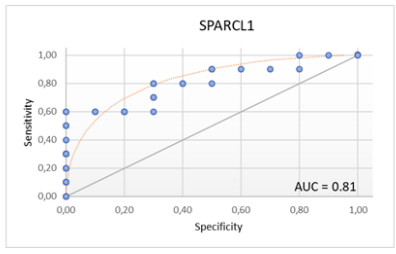

VCL
Sensitivity
Specificity
AUC = 0.79

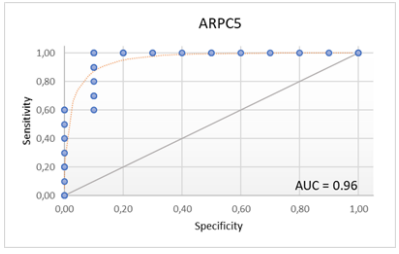

RPL7
Sensitivity
Specificity
AUC = 0.97


TUFM
Sensitivity
Specificity
AUC = 0.90


HMGB2
Sensitivity
Specificity
AUC = 0.82


CDH1
Sensitivity
Specificity
AUC = 0.70


SPARCL1
Sensitivity
Specificity
AUC = 0.81


ARPC5
Sensitivity
Specificity
AUC = 0.96

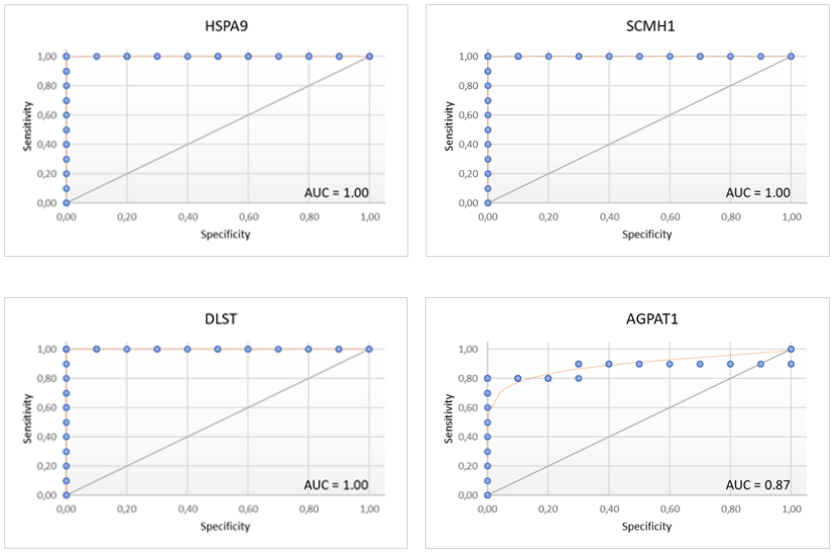


Figure 58. ROC analysis of proteins with significantly (**A**) higher and (**B**) lower expression in the saliva of patients at early OSCC stages collated to PVL. AUC values indicate the capacity of each marker to discriminate between patients with PVL and early OSCC lesions.

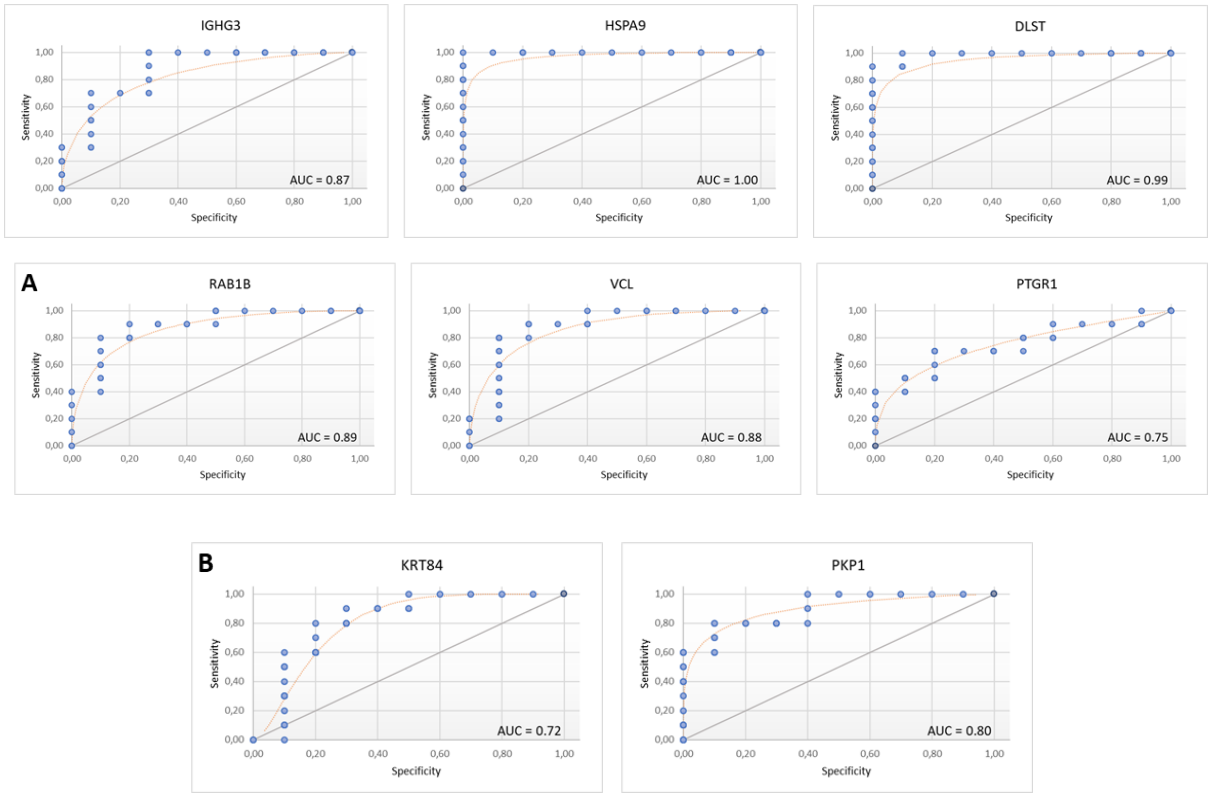


Figure 61. ROC curves and corresponding AUC statistics for salivary proteins with notably (**A**) elevated and (**B**) reduced relative abundance in advanced OSCC compared to PVL patients.

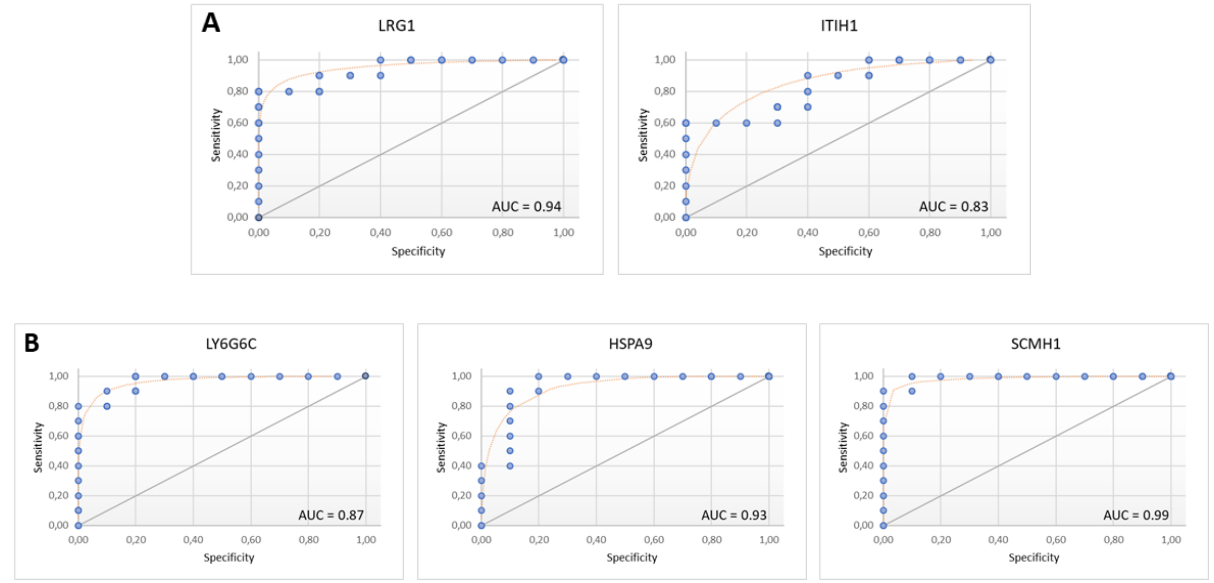
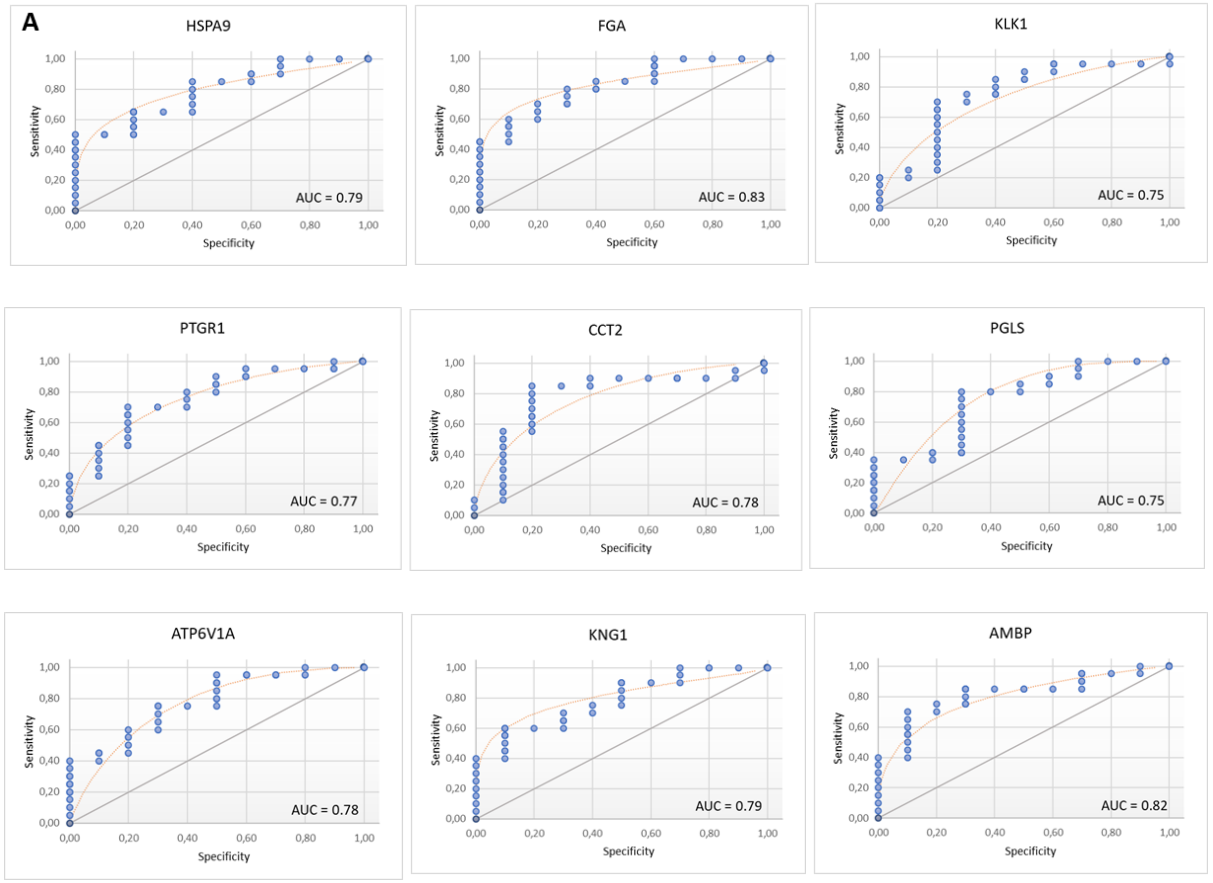


Figure 63. ROC curves and corresponding area under the curve (AUC) values for importantly (**A**) upregulated and (**B**) downregulated salivary proteins of OSCC collated to PVL patients. The area under the ROC curve is the measure of predictive models to discern between PVL and OSCC patients.



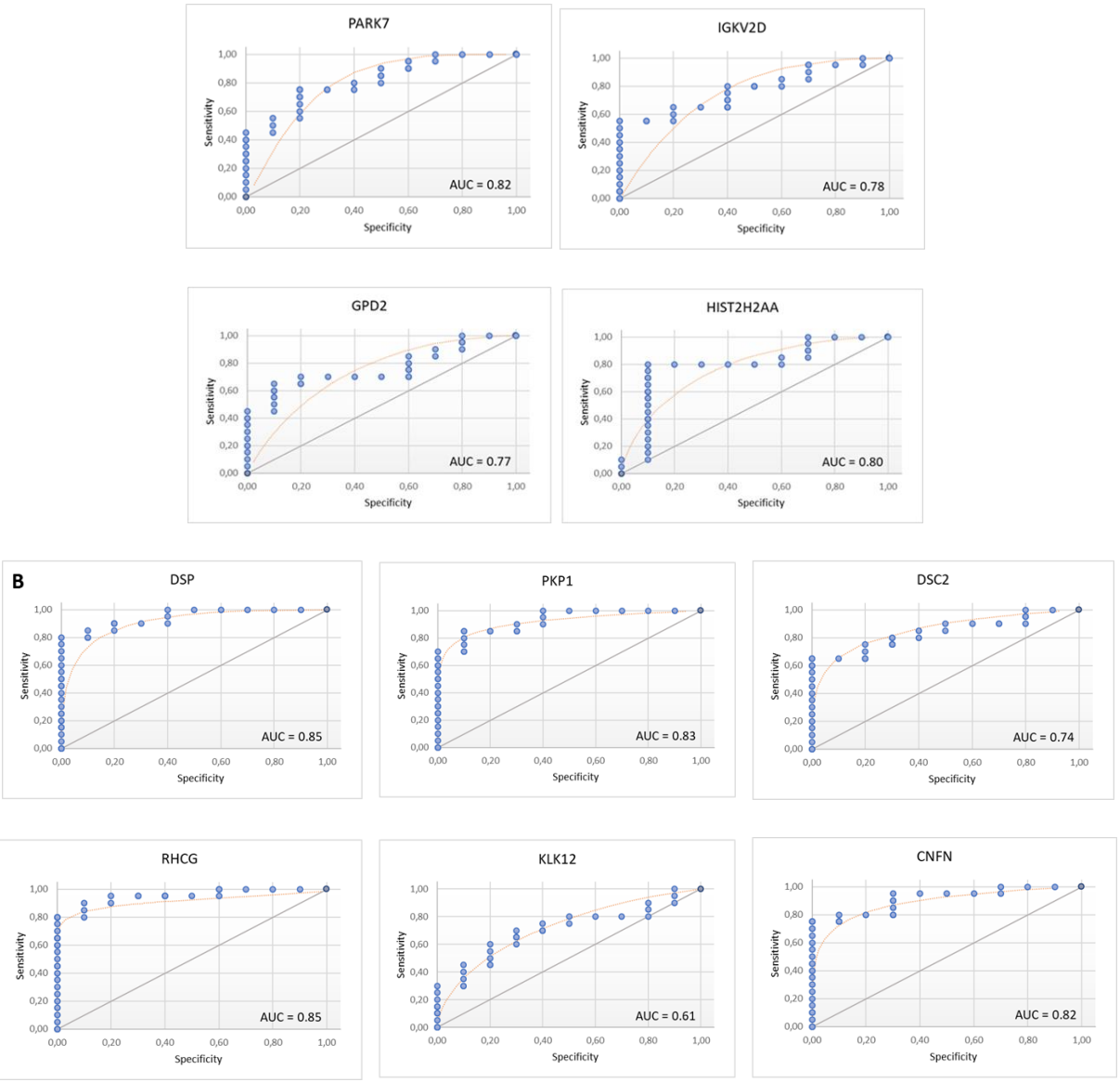
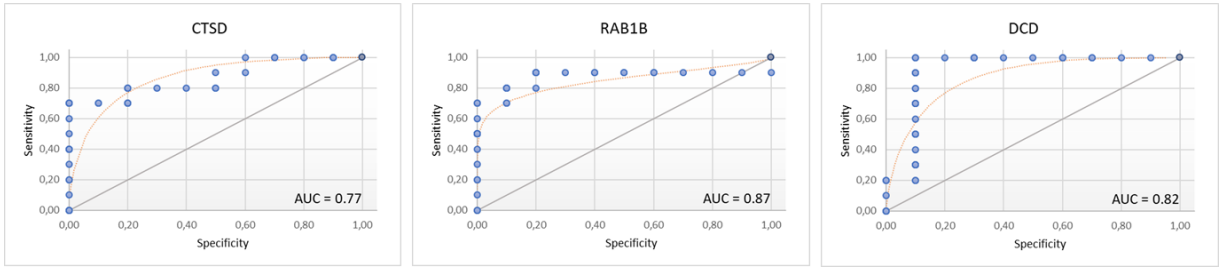
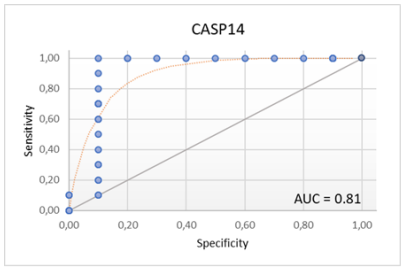
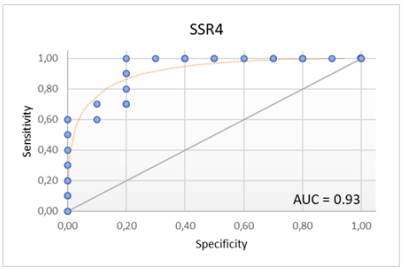
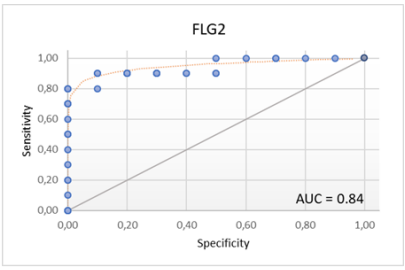


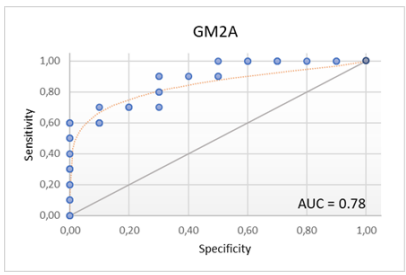
Figure 66. ROC analysis and corresponding AUC statistics for upregulated proteins in early compared to advanced OSCC saliva. The area under the ROC curve is the measure of predictive models to discern between early and advanced disease stages.

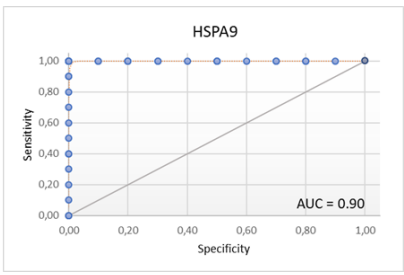


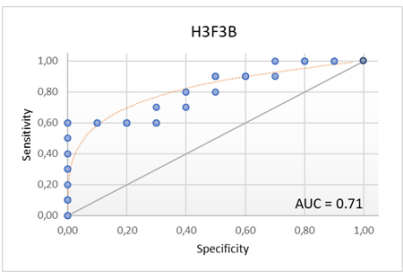

CASP14
Sensitivity
Specificity
AUC = 0.81

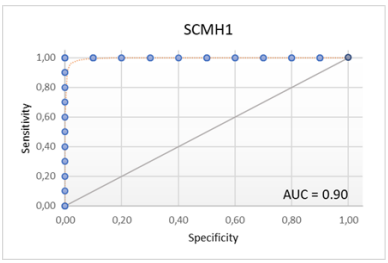

SSR4
Sensitivity
Specificity
AUC = 0.93

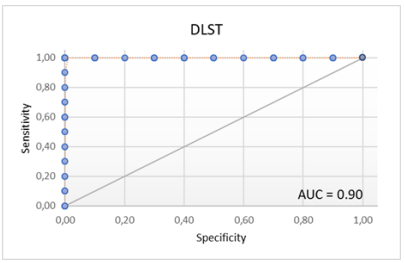

FLG2
Sensitivity
Specificity
AUC = 0.84


GM2A
Sensitivity
Specificity
AUC = 0.78


HSPA9
Sensitivity
Specificity
AUC = 0.90


H3F3B
Sensitivity
Specificity
AUC = 0.71


SCMH1
Sensitivity
Specificity
AUC = 0.90


DLST
Sensitivity
Specificity
AUC = 0.90

X. *SUPPLEMENTARY TABLES*

Table 6. Variability measures of cytokine expression values in control, HL, PVL, early and advanced OSCC groups.

Control	1st Qu	3rd Qu	Median	Mean	Min	Max	Variance	SD	CV
IL-1α	807,30	1611,50	1268,40	1227,20	217,40	2264,80	377611,20	614,50	0,501
IL-6	5,38	9,35	7,40	7,95	1,13	21,37	22,63	4,76	0,599
IL-8	318,70	716,60	420,90	526,20	151,50	1262,70	87123,69	295,17	0,561
IP-10	592,60	1137,60	756,00	885,00	320,40	1927,20	218035,00	466,94	0,528
MCP-1	503,20	1586,70	977,10	1066,60	326,90	2343,20	400328,90	632,72	0,593
TNF-α	4,26	10,81	7,12	7,62	1,59	17,00	17,64	4,20	0,551
HCC-1	58,76	82,71	70,48	75,36	21,90	149,35	930,69	30,51	0,405
PF-4	175,00	296,00	231,00	253,80	89,00	544,00	14373,44	119,89	0,472
HL	1st Qu	3rd Qu	Median	Mean	Min	Max	Variance	SD	CV
IL-1α	558,00	1701,50	1042,50	1760,10	114,60	8986,00	4838564,00	2199,67	1,250
IL-6	6,72	20,93	11,69	22,45	2,71	121,22	778,54	27,90	1,243
IL-8	367,46	1541,77	947,25	1405,26	48,52	7778,78	2645994,00	1626,65	1,158
IP-10	178,55	2531,63	759,91	1695,40	27,45	9340,92	5017955,00	2240,08	1,321
MCP-1	1047,90	3779,00	1694,30	2588,40	303,60	8145,40	4639360,00	2153,92	0,832
TNF-α	7,12	27,08	14,38	19,88	1,27	58,09	268,31	16,38	0,824
HCC-1	69,24	221,96	99,45	175,46	23,26	755,45	26041,28	161,37	0,920
PF-4	116,80	325,50	188,00	300,90	24,00	1684,00	123166,20	350,95	1,166
PVL	1st Qu	3rd Qu	Median	Mean	Min	Max	Variance	SD	CV
IL-1α	640,42	2450,29	1479,68	1886,48	18,52	7527,01	2869304,00	1693,90	0,898
IL-6	7,32	18,62	12,77	18,90	3,03	91,99	464,23	21,55	1,140
IL-8	479,80	1280,80	741,90	1140,90	102,30	7954,20	1908385,00	1381,44	1,211
IP-10	247,33	1834,62	541,03	1157,80	40,47	5577,07	2084033,00	1443,62	1,247
MCP-1	1489,90	5608,30	2904,70	3600,80	418,90	9218,80	7645029,00	2764,96	0,768
TNF-α	9,00	29,57	14,49	23,08	2,03	118,69	599,03	24,48	1,061
HCC-1	74,07	182,12	126,15	189,28	26,96	1042,43	42213,02	205,46	1,085
PF-4	97,00	372,00	175,00	454,80	10,00	3793,00	709393,50	842,26	1,852
early OSCC	1st Qu	3rd Qu	Median	Mean	Min	Max	Variance	SD	CV
IL-1α	574,95	2469,64	1053,16	1556,19	46,94	6082,49	2034714,00	1426,43	0,917
IL-6	14,57	117,87	52,56	99,82	3,00	747,10	22457,29	149,86	1,501
IL-8	761,80	2495,50	1413,70	2567,00	114,50	11491,60	9972508,00	3157,93	1,230
IP-10	592,50	2169,70	1132,60	1963,90	131,50	7156,00	4406245,00	2099,11	1,069
MCP-1	1036,60	4304,20	1582,20	2560,70	380,20	7603,40	4524078,00	2126,99	0,831
TNF-α	25,09	65,43	37,35	59,82	4,88	310,11	4412,39	66,43	1,110
HCC-1	131,89	374,11	210,96	298,53	27,33	1511,40	81008,72	284,62	0,953
PF-4	326,00	727,00	513,00	642,10	33,00	3069,00	362741,50	602,28	0,938
advanced OSCC	1st Qu	3rd Qu	Median	Mean	Min	Max	Variance	SD	CV
IL-1α	436,30	1744,80	884,10	1313,80	159,00	3961,10	1186101,00	1089,08	0,829
IL-6	60,80	358,73	119,46	262,08	7,15	1711,81	143005,70	378,16	1,443
IL-8	1218,41	5544,88	2617,79	4124,81	61,09	16314,10	20451026,00	4522,28	1,096
IP-10	456,70	1880,30	631,30	1556,50	65,60	7673,40	3646675,00	1909,63	1,227
MCP-1	732,90	1905,40	1184,00	2099,80	119,10	9900,70	6100394,00	2469,90	1,176
TNF-α	34,46	160,48	91,88	122,52	6,03	504,32	14113,07	118,80	0,970
HCC-1	177,00	600,50	270,00	551,30	32,00	3404,60	457948,20	676,72	1,227
PF-4	190,50	1713,50	306,20	1021,90	115,00	4658,00	1566097,00	1251,44	1,225

1st Qu and **3rd Q** - first and third quartile, respectively; SD- standard deviation; **CV**-coefficient of variation

Table 8. Pearson's pairwise correlation among salivary cytokines in (**A**) Control, (**B**) HL, (**C**) PVL, (**D**) early OSCC, and (**E**) advanced OSCC group. Correlation is significant when p ≤ 0.05 (*), p ≤ 0.01 (**) and, p ≤ 0.001 (***).

A		**IL-8**	**MCP-1**	**TNF-α**	**HCC-1**	**PF-4**
IL-α	Pearson correlation	* 0,426		* 0,458		
	Sig. (two-tailed)	0,03		0,02		
IL-6	Pearson correlation	** 0,593	* 0,493	*** 0,817	*0,438	
	Sig. (two-tailed)	0,002	0,01	< 0,001	0,03	
IL-8	Pearson correlation			*** 0,706	*0,512	
	Sig. (two-tailed)			< 0,001	0,01	
TNF-α	Pearson correlation				*0,463	
	Sig. (two-tailed)				0,02	
HCC-1	Pearson correlation					** 0,601
	Sig. (two-tailed)					0,001

B		**IL-8**	**IP-10**	**MCP-1**	**TNF-α**	**HCC-1**	**PF-4**
IL-6	Pearson correlation	** 0,459			***0,781	***0,667	*0,383
	Sig. (two-tailed)	0,007			< 0,001	< 0,001	0,03
IL-8	Pearson correlation		**0,446	**0,464	***0,598	**0,475	*0,388
	Sig. (two-tailed)		0,009	0,007	< 0,001	0,01	0,03
IP-10	Pearson correlation					***0,667	**0,466
	Sig. (two-tailed)					< 0,001	0,006
MCP-1	Pearson correlation				**0,507		
	Sig. (two-tailed)				0,003		
TNF-α	Pearson correlation					***0,604	
	Sig. (two-tailed)					< 0,001	
HCC-1	Pearson correlation						**0,499
	Sig. (two-tailed)						0,003

C		**IL-6**	**IL-8**	**IP-10**	**MCP-1**	**HCC-1**	**PF-4**
IL-1α	Pearson correlation	**0,522	***0,662		*0,375	**0,496	
	Sig. (two-tailed)	0,002	< 0,001		0,03	0,003	
IL-6	Pearson correlation		**0,492		*0,406	**0,478	*0,436
	Sig. (two-tailed)		0,004		0,02	0,005	0,01
IL-8	Pearson correlation			*0,362	*0,372	***0,754	
	Sig. (two-tailed)			0,04	0,03	< 0,001	
IP-10	Pearson correlation				*0,392	*0,382	
	Sig. (two-tailed)				0,02	0,03	
MCP-1	Pearson correlation					***0,559	***0,597
	Sig. (two-tailed)					< 0,001	< 0,001
TNF-α	Pearson correlation					**0,453	*0,411
	Sig. (two-tailed)					0,008	0,02

D		IL-8	MCP-1	TNF-α	HCC-1	PF-4
IL-1α	Pearson correlation		*0,364	**0,528		
	Sig. (two-tailed)		0,04	0,002		
IL-6	Pearson correlation	**0,478		**0,456		
	Sig. (two-tailed)	0,005		0,008		
IL-8	Pearson correlation		**0,515	***0,619	**0,522	*0,350
	Sig. (two-tailed)		0,002	< 0,001	0,002	0,05
MCP-1	Pearson correlation			*0,393		
	Sig. (two-tailed)			0,02		
TNF-α	Pearson correlation				***0,595	
	Sig. (two-tailed)				< 0,001	

E		IP-10	MCP-1	TNF-α	HCC-1	PF-4
IL-6	Pearson correlation	**0,457		***0,568		
	Sig. (two-tailed)	0,007		< 0,001		
IL-8	Pearson correlation		***0,549		***0,558	***0,738
	Sig. (two-tailed)		< 0,001		< 0,001	< 0,001
MCP-1	Pearson correlation				***0,569	**0,495
	Sig. (two-tailed)				< 0,001	0,003
HCC-1	Pearson correlation					*0,369
	Sig. (two-tailed)					0,03

Table 9. Conditional concentration ranges (pg/mL) of salivary cytokines that maximize the probability of association with the clinical variables (**A**) histologic feature – epithelial dysplasia, (**B**) histologic feature - no epithelial dysplasia, (**C**) clinical type- mixed, (**D**) clinical type – verrucous and (**E**) clinical type homogeneous white lesion of patients diagnosed with HL and PVL. **Mean**- average value (n=33 HL; n=33 PVL); **L**- lower and **U**- upper value; **P**- probability of diagnosis. Cytokines, positively correlated with patients' clinical variables are marked in red.

A	Histologic feature - Epithelial Dysplasia (ED)				
	mean	L	U	P (ED)	P (no ED)
IL-1α	2140,72	2120,73	2160,72	0.495 [+-0.0662]	0.513 [+-0.0644]
IL-6	20,2897	20,16	20,40	0.52 [+-0.0775]	0.486 [+-0.072]
IL-8	1337,36	1325,85	1348,86	0.491 [+-0.0649]	0.509 [+-0.0619]
IP-10	1769,37	1739,23	1799,51	0.499 [+-0.062]	0.508 [+-0.0644]
MCP-1	3090,77	3072,57	3108,97	0.497 [+-0.0725]	0.507 [+-0.0755]
TNF-α	22,77	22,58	22,96	0.513 [+-0.0612]	0.491 [+-0.0696]
HCC-1	168,84	168,06	169,62	0.53 [+-0.0778]	0.47 [+-0.0662]
PF-4	375,22	371,36	379,07	0.497 [+-0.0644]	0.507 [+-0.0737]

B	Histologic feature - No Epithelial Dysplasia (no ED)				
	mean	L	U	P (no ED)	P (ED)
IL-1α	2107,27	2088,40	2126,13	0.512 [+-0.0684]	0.496 [+-0.0688]
IL-6	22,20	22,05	22,36	0.492 [+-0.0622]	0.503 [+-0.0656]
IL-8	1338,61	1327,30	1349,93	0.513 [+-0.0634]	0.497 [+-0.0601]
IP-10	2164,79	2117,94	2211,64	0.507 [+-0.0624]	0.5 [+-0.0529]
MCP-1	3276,44	3256,06	3296,83	0.503 [+-0.0704]	0.502 [+-0.0724]
TNF-α	25,41	25,17	25,65	0.494 [+-0.069]	0.515 [+-0.0583]
HCC-1	199,78	198,46	201,09	0.479 [+-0.0736]	0.518 [+-0.0677]
PF-4	376,45	372,64	380,27	0.485 [+-0.0683]	0.489 [+-0.0673]

C	Clinical type - Mixed					
	mean	L	U	P (Mixed)	P (HWL)	P (Verrucous)
IL-1α	1901,22	1874,24	1928,20	0.205 [+-0.0414]	0.511 [+-0.0548]	0.285 [+-0.0474]
IL-6	16,32	16,18	16,47	0.204 [+-0.0479]	0.485 [+-0.0537]	0.308 [+-0.0515]
IL-8	1218,48	1205,55	1231,41	0.189 [+-0.0412]	0.42 [+-0.0507]	0.38 [+-0.0492]
IP-10	950,68	934,24	967,12	0.212 [+-0.0404]	0.479 [+-0.0525]	0.304 [+-0.0479]
MCP-1	2955,81	2927,85	2983,76	0.21 [+-0.0507]	0.5 [+-0.056]	0.293 [+-0.043]
TNF-α	16,08	15,96	16,20	0.268 [+-0.0634]	0.5 [+-0.0692]	0.235 [+-0.0576]
HCC-1	169,52	168,07	170,96	0.198 [+-0.0391]	0.486 [+-0.055]	0.32 [+-0.0534]
PF-4	272,73	268,05	277,41	0.159 [+-0.0337]	0.577 [+-0.0503]	0.268 [+-0.0393]

D	Clinical type - Verrucous					
	mean	L	U	P (Verrucous)	P (HWP)	P(Mixed)
IL-1α	2357,65	2327,98	2387,33	0.298 [+-0.0576]	0.503 [+-0.0627]	0.198 [+-0.0414]
IL-6	21,09	20,93	21,24	0.311 [+-0.0671]	0.503 [+-0.0668]	0.187 [+-0.0488]
IL-8	962,06	957,53	966,60	0.407 [+-0.074]	0.407 [+-0.0762]	0.184 [+-0.0547]
IP-10	1616,98	1585,52	1648,43	0.297 [+-0.0488]	0.515 [+-0.0628]	0.185 [+-0.0391]
MCP-1	3277,21	3252,16	3302,26	0.291 [+-0.0577]	0.499 [+-0.0593]	0.204 [+-0.0515]
TNF-α	30,25	29,83	30,67	0.269 [+-0.0502]	0.539 [+-0.0517]	0.196 [+-0.0422]
HCC-1	175,44	174,41	176,48	0.319 [+-0.0621]	0.484 [+-0.0647]	0.204 [+-0.0569]
PF-4	607,86	599,77	615,95	0.374 [+-0.067]	0.473 [+-0.0674]	0.145 [+-0.0461]

E	Clinical type - Homogeneous While Plaques (HWP)					
	mean	L	U	P (HWP)	P (Mixed)	P (Verrucous)
IL-1α	2090,17	2071,80	2108,54	0.521 [+-0.0667]	0.193 [+-0.0567]	0.291 [+-0.0592]
IL-6	23,43	23,27	23,59	0.514 [+-0.0674]	0.18 [+-0.0542]	0.32 [+-0.0632]
IL-8	1592,65	1578,13	1607,18	0.487 [+-0.0691]	0.204 [+-0.0515]	0.302 [+-0.065]
IP-10	2590,62	2538,99	2642,25	0.549 [+-0.0583]	0.153 [+-0.0455]	0.292 [+-0.0545]
MCP-1	3231,22	3211,25	3251,19	0.497 [+-0.0725]	0.2 [+-0.0616]	0.299 [+-0.0669]
TNF-α	24,15	23,96	24,34	0.521 [+-0.0693]	0.229 [+-0.0561]	0.26 [+-0.0578]
HCC-1	195,80	194,53	197,07	0.491 [+-0.0665]	0.201 [+-0.0513]	0.325 [+-0.0638]
PF-4	293,08	291,09	295,06	0.577 [+-0.0751]	0.157 [+-0.0516]	0.281 [+-0.0676]

Table 10. Conditional concentration ranges (pg/mL) of salivary cytokines that maximize the probability of association with the clinical variables (**A**) histologic feature – well-differentiated tumor (WD), (**B**) histologic feature – not well-differentiated tumor, (**C**) absence of cervical adenopathy, and (**D**) the presence of cervical adenopathy of patients diagnosed at early and advanced OSCC stages. **Mean**- average value; **L**- lower and **U**- upper value; **P**- the probability of association. Analytes, positively correlated with patients' clinical variables are marked in red.

A	Histologic features - well differentiated (WD)				
	mean	L	U	P (WD)	P (not WD)
IL-1α	1707,16	1696,67	1717,66	0.734 [+-0.0601]	0.14 [+-0.0536]
IL-6	176,61	174,09	179,13	0.687 [+-0.0596]	0.178 [+-0.0509]
IL-8	7,50	7,49	7,50	0 [+-0]	0 [+-0]
IP-10	1882,14	1868,03	1896,25	0.731 [+-0.0652]	0.146 [+-0.0567]
MCP-1	2966,31	2949,49	2983,13	0.838 [+-0.0544]	0.057 [+-0.0407]
TNF-α	88,27	87,62	88,93	0.715 [+-0.0659]	0.166 [+-0.0541]
HCC-1	364,69	362,87	366,51	0.751 [+-0.0703]	0.153 [+-0.0569]
PF-4	899,59	893,06	906,12	0.739 [+-0.0655]	0.143 [+-0.0479]

B	Histologic features - not well differentiated (WD)				
	mean	L	U	P (not WD)	P (WD)
IL-1α	1691,62	1668,39	1714,86	0.137 [+-0.0344]	0.744 [+-0.0459]
IL-6	352,30	340,73	363,88	0.217 [+-0.0467]	0.643 [+-0.0512]
IL-8	7,73	7,71	7,74	0.007 [+-0.0816]	0 [+-0]
IP-10	2244,84	2201,22	2288,46	0.157 [+-0.0339]	0.727 [+-0.0466]
MCP-1	1076,91	1065,31	1088,51	0.163 [+-0.0385]	0.717 [+-0.0527]
TNF-α	174,59	172,08	177,10	0.233 [+-0.0543]	0.618 [+-0.0658]
HCC-1	737,30	729,15	745,45	0.242 [+-0.0549]	0.625 [+-0.0628]
PF-4	1002,83	986,16	1019,50	0.144 [+-0.0341]	0.726 [+-0.0425]

C	No cervical adenopathy (CA)				
	mean	L	U	P (no CA)	P (CA)
IL-1α	1706,70	1692,79	1720,61	0.423 [+-0.0608]	0.58 [+-0.0623]
IL-6	72,28	71,62	72,93	0.478 [+-0.0767]	0.536 [+-0.0625]
IL-8	7,22	7,21	7,22	0 [+-0]	0 [+-0]
IP-10	1664,28	1648,14	1680,42	0.409 [+-0.0592]	0.592 [+-0.0666]
MCP-1	2375,25	2358,58	2391,92	0.433 [+-0.0675]	0.56 [+-0.0624]
TNF-α	64,24	63,68	64,81	0.427 [+-0.0645]	0.581 [+-0.061]
HCC-1	358,66	355,98	361,35	0.413 [+-0.0652]	0.591 [+-0.0594]
PF-4	773,68	766,61	780,74	0.409 [+-0.0663]	0.584 [+-0.0673]

D	Cervical adenopathy (CA)				
	mean	L	U	P (CA)	P (no CA)
IL-1α	1706,23	1694,39	1718,07	0.582 [+-0.0668]	0.421 [+-0.0662]
IL-6	313,46	308,94	317,97	0.817 [+-0.0629]	0.181 [+-0.0649]
IL-8	7,77	7,77	7,78	0.007 [+-0.0816]	0 [+-0]
IP-10	2171,82	2152,56	2191,08	0.606 [+-0.0683]	0.387 [+-0.0707]
MCP-1	2884,44	2864,19	2904,69	0.569 [+-0.0744]	0.423 [+-0.0694]
TNF-α	136,34	135,26	137,43	0.732 [+-0.0735]	0.275 [+-0.0718]
HCC-1	559,70	555,17	564,23	0.634 [+-0.0675]	0.355 [+-0.0635]
PF-4	1024,76	1016,33	1033,18	0.607 [+-0.0684]	0.384 [+-0.0659]

Table 11. Conditional probability ranges of salivary cytokines to distinguish between (**A**) Control, (**B**) HL, (**C**) PVL, (**D**) early, and (**D**) advanced OSCC stages. **Mean**- average value; **L**- lower and **U**- upper value; **P**- the probability of association. Cytokines, positively correlated with patients' clinical variables are marked in red.

A	Control							
	mean	L	U	P (Control)	P (HL)	P (advanced OSCC stage)	P (early OSCC stages)	P (PVL)
IL-α	1270,75	1255,78	1285,72	0.171 [+-0.0403]	0.246 [+-0.0471]	0.187 [+-0.0419]	0.171 [+-0.0408]	0.233 [+-0.0485]
IL-6	8,43	8,36	8,49	0.289 [+-0.0781]	0.282 [+-0.0587]	0.025 [+-0.0226]	0.085 [+-0.0457]	0.316 [+-0.0709]
IL-8	536,26	533,08	539,43	0.332 [+-0.0727]	0.23 [+-0.0685]	0.088 [+-0.0422]	0.113 [+-0.0469]	0.267 [+-0.0702]
IP-10	894,90	890,19	899,62	0.353 [+-0.0775]	0.155 [+-0.0632]	0.165 [+-0.0581]	0.158 [+-0.0578]	0.174 [+-0.0581]
MCP-1	1099,06	1091,56	1106,55	0.265 [+-0.063]	0.243 [+-0.0596]	0.179 [+-0.0491]	0.163 [+-0.0507]	0.164 [+-0.0472]
TNF-α	7,89	7,84	7,94	0.376 [+-0.0779]	0.288 [+-0.0783]	0.017 [+-0.0229]	0.059 [+-0.0393]	0.252 [+-0.0703]
HCC-1	76,72	76,38	77,06	0.397 [+-0.0841]	0.246 [+-0.0743]	0.067 [+-0.0414]	0.078 [+-0.048]	0.223 [+-0.0664]
PF-4	256,32	255,11	257,52	0.352 [+-0.0715]	0.217 [+-0.0658]	0.121 [+-0.0506]	0.142 [+-0.059]	0.155 [+-0.0562]

B	Homogeneous Leukoplakia (HL)							
	mean	L	U	P (HL)	P (Control)	P (advanced OSCC stages)	P (early OSCC stages)	P (PVL)
IL-α	1833,92	1813,27	1854,58	0.247 [+-0.0486]	0.147 [+-0.0433]	0.175 [+-0.0423]	0.176 [+-0.0417]	0.249 [+-0.0442]
IL-6	21,98	21,77	22,20	0.328 [+-0.0625]	0.091 [+-0.0379]	0.093 [+-0.0418]	0.178 [+-0.0555]	0.329 [+-0.0607]
IL-8	1507,48	1488,64	1526,33	0.261 [+-0.0585]	0.051 [+-0.0267]	0.187 [+-0.0497]	0.209 [+-0.0495]	0.295 [+-0.0565]
IP-10	2332,65	2269,57	2395,73	0.218 [+-0.0439]	0.081 [+-0.0294]	0.233 [+-0.0407]	0.262 [+-0.0471]	0.209 [+-0.0382]
MCP-1	2612,81	2591,56	2634,07	0.281 [+-0.0574]	0.077 [+-0.0361]	0.166 [+-0.0508]	0.204 [+-0.0573]	0.27 [+-0.0585]
TNF-α	21,61	21,39	21,83	0.308 [+-0.0582]	0.06 [+-0.0322]	0.097 [+-0.0378]	0.224 [+-0.0552]	0.309 [+-0.0586]
HCC-1	179,01	177,52	180,51	0.278 [+-0.0612]	0.043 [+-0.0273]	0.159 [+-0.0491]	0.218 [+-0.0565]	0.296 [+-0.0566]
PF-4	296,13	293,06	299,20	0.217 [+-0.0445]	0.326 [+-0.058]	0.13 [+-0.0384]	0.167 [+-0.0405]	0.157 [+-0.0417]

C	Proliferative Verrucous Leukoplakia (PVL)							
	mean	L	U	P (PVL)	P (Control)	P (HL)	P (advanced OSCC stages)	P (early OSCC stages)
IL-α	2163,32	2139,04	2187,60	0.255 [+-0.0542]	0.141 [+-0.0397]	0.245 [+-0.0553]	0.185 [+-0.0509]	0.184 [+-0.0479]
IL-6	18,77	18,61	18,93	0.333 [+-0.0659]	0.106 [+-0.0452]	0.32 [+-0.0741]	0.072 [+-0.0355]	0.15 [+-0.053]
IL-8	1160,30	1150,06	1170,55	0.304 [+-0.0591]	0.099 [+-0.0418]	0.259 [+-0.0611]	0.148 [+-0.0465]	0.183 [+-0.0567]
IP-10	1389,10	1363,17	1415,04	0.184 [+-0.0362]	0.256 [+-0.0431]	0.172 [+-0.0376]	0.195 [+-0.039]	0.202 [+-0.0417]
MCP-1	3902,11	3864,80	3939,42	0.32 [+-0.069]	0.03 [+-0.0224]	0.277 [+-0.065]	0.166 [+-0.0544]	0.212 [+-0.0549]
TNF-α	24,71	24,46	24,97	0.309 [+-0.0574]	0.046 [+-0.0306]	0.29 [+-0.0566]	0.115 [+-0.0404]	0.256 [+-0.0557]
HCC-1	191,01	189,48	192,53	0.3 [+-0.0622]	0.038 [+-0.0276]	0.283 [+-0.0625]	0.163 [+-0.0448]	0.238 [+-0.057]
PF-4	477,63	469,33	485,94	0.184 [+-0.0349]	0.158 [+-0.0363]	0.209 [+-0.0455]	0.198 [+-0.0437]	0.257 [+-0.0429]

D	early OSCC stages							
	mean	L	U	P (early OSCC stages)	P (Control)	P (HL)	P (advanced OSCC stages)	P (PVL)
IL-α	1906,48	1881,49	1931,46	0.181 [+-0.04]	0.151 [+-0.0439]	0.247 [+-0.0438]	0.172 [+-0.0425]	0.25 [+-0.0472]
IL-6	82,73	81,15	84,31	0.363 [+-0.0744]	0.001 [+-0.0035]	0.132 [+-0.0465]	0.43 [+-0.0682]	0.078 [+-0.0409]
IL-8	3236,26	3177,11	3295,40	0.278 [+-0.061]	0.003 [+-0.0064]	0.234 [+-0.0538]	0.316 [+-0.0584]	0.167 [+-0.0427]
IP-10	2095,11	2062,70	2127,52	0.244 [+-0.0518]	0.112 [+-0.0372]	0.205 [+-0.0517]	0.237 [+-0.0522]	0.203 [+-0.0435]
MCP-1	2848,00	2819,54	2876,46	0.213 [+-0.0561]	0.062 [+-0.0303]	0.282 [+-0.0594]	0.164 [+-0.0479]	0.274 [+-0.0554]
TNF-α	54,82	54,26	55,37	0.332 [+-0.0827]	0.002 [+-0.0066]	0.156 [+-0.0579]	0.307 [+-0.0746]	0.198 [+-0.0603]
HCC-1	344,64	341,48	347,81	0.297 [+-0.0694]	0.001 [+-0.0035]	0.203 [+-0.0571]	0.251 [+-0.0592]	0.229 [+-0.0571]
PF-4	688,69	681,07	696,31	0.316 [+-0.0712]	0.048 [+-0.0286]	0.184 [+-0.0614]	0.254 [+-0.0624]	0.195 [+-0.058]

E	advanced OSCC stages							
	mean	L	U	P (advanced OSCC stages)	P (Control)	P (HL)	P (OSCC early stages)	P (PVL)
IL-α	1698,81	1677,45	1720,16	0.181 [+-0.0384]	0.155 [+-0.0396]	0.244 [+-0.0446]	0.174 [+-0.0379]	0.247 [+-0.0498]
IL-6	272,35	267,50	277,20	0.723 [+-0.1061]	0 [+-0]	0.013 [+-0.0217]	0.256 [+-0.0902]	0.002 [+-0.0103]
IL-8	6097,95	5963,90	6232,01	0.447 [+-0.0727]	0 [+-0]	0.171 [+-0.0634]	0.313 [+-0.0738]	0.07 [+-0.0413]
IP-10	1879,83	1848,76	1910,91	0.218 [+-0.0493]	0.139 [+-0.0368]	0.197 [+-0.0489]	0.242 [+-0.0485]	0.191 [+-0.0459]
MCP-1	2422,93	2394,01	2451,86	0.173 [+-0.0383]	0.087 [+-0.028]	0.285 [+-0.0439]	0.201 [+-0.0414]	0.256 [+-0.0472]
TNF-α	156,08	154,10	158,05	0.654 [+-0.1021]	0 [+-0]	0.046 [+-0.0428]	0.234 [+-0.0917]	0.064 [+-0.0492]
HCC-1	632,54	623,03	642,04	0.407 [+-0.0771]	0 [+-0.0015]	0.133 [+-0.0573]	0.327 [+-0.0748]	0.152 [+-0.0575]
PF-4	1257,39	1233,62	1281,15	0.356 [+-0.0679]	0.003 [+-0.0076]	0.114 [+-0.0487]	0.316 [+-0.0607]	0.207 [+-0.0586]

Table 13. Annotation of the total *N*-glycans identified in the salivary pool of healthy individuals on HILIC-UPLC–MS identification. **dHex** –desoxyhexose (fucose), **Hex**- hexose (galactose/glucose/mannose), **HexNAc**- N-acetylgalactosamine/N-acetylglucosamine, **NeuAc** – sialic acid, **Pen**-Xylose. *N*-glycans found to have significantly different relative abundance in OSCC compared to PVL and Controls groups are highlighted.

$t_{(R)}$ min	Experimental mass (m/z)	Theoretical Monoisotopic mass (m/z)	Mass accuracy (ppm)	Ion	Monosacharide composition
3.86	1175.4491	1175.4472	1.65	$[M-H]^{1-}$	dHex-core2AB
4.99	688.7597	688.7596	0.10	$[M2-H]^{2-}$	dHexHexNAc-core2AB
5.20	717.2722	717.2704	2.60	$[M2-H]^{2-}$	HexNAc2-core2AB
5.58	668.2479	668.2464	2.31	$[M2-H]^{2-}$	dHexHex-core2AB
	1337.5014	1337.5	1.05	$[M-H]^{1-}$	dHexHex-core2AB
5.90	696.7568	696.7571	0.41	$[M2-H]^{2-}$	HexHexNAc-core2AB
6.19	818.8082	818.81	2.25	$[M2-H]^{2-}$	HexNAc3-core2AB
6.27	790.2980	790.2993	1.66	$[M2-H]^{2-}$	dHexHexNAc2-core2AB
6.88	676.2428	676.2438	1.49	$[M2-H]^{2-}$	Hex2-core2AB
7.24	769.7853	769.786	0.96	$[M2-H]^{2-}$	dHexHexHexNAc-core2AB
7.24	891.8372	891.839	2.02	$[M2-H]^{2-}$	dHexHexNAc3-core2AB
8.14	899.8359	899.8365	0.62	$[M2-H]^{2-}$	HexHexNAc3-core2AB
8.25	842.8142	842.8150	0.94	$[M2-H]^{2-}$	dHex2HexNAcHex-core2AB
8.68	871.3248	871.3257	1.06	$[M2-H]^{2-}$	dHexHexNAc2Hex-core2AB
9.24	972.8630	972.8654	2.47	$[M2-H]^{2-}$	dHexHexNAc3Hex-core2AB
9.37	757.2692	757.2702	1.35	$[M2-H]^{2-}$	Hex3-core2AB
9.46	850.8111	850.8125	1.58	$[M2-H]^{2-}$	dHexHexNAcHex2-core2AB
9.53	879.3229	879.3232	0.32	$[M2-H]^{2-}$	Hex2HexNAc2-core2AB
10.01	915.3345	915.3337	0.82	$[M2-H]^{2-}$	dHexNeuAcHexNAcHex-core2AB
10.27	915.8436	915.8439	0.38	$[M2-H]^{2-}$	dHex3HexNAcHex-core2AB
10.61	944.3545	944.3547	0.19	$[M2-H]^{2-}$	dHex2HexNAc2Hex-core2AB
10.93	952.3506	952.3521	1.61	$[M2-H]^{2-}$	dHex1HexNAc2Hex2-core2AB
11.34	915.3335	915.3335	0.27	$[M2-H]^{2-}$	dHexNeuAcHexNAcHex-core2AB
11.59	923.8405	923.8414	0.99	$[M2-H]^{2-}$	dHex2HexNAcHex2-core2AB
11.62	1053.8900	1053.8918	1.73	$[M2-H]^{2-}$	dHexHexNAc3Hex2-core2AB
11.72	1045.8924	1045.8944	1.88	$[M2-H]^{2-}$	dHex2HexNAc3Hex-core2AB

11.97	1025.3796	1025.3811	1.45	$[M2-H]^{2-}$	dHex2HexNAc2Hex2-core2AB
12.05	1017.3822	1017.3836	1.41	$[M2-H]^{2-}$	dHex3HexNAc2Hex2-core2AB
12.17	988.3617	988.3627	1.01	$[M2-H]^{2-}$	dHex2NeuAcHexNAcHex-core2AB
12.19	931.8380	931.8389	0.93	$[M2-H]^{2-}$	dHexHexNAcHex3-core2AB
12.46	996.3601	996.3602	0.06	$[M2-H]^{2-}$	dHexNeuAcHexNAcHex2-core2AB
12.72	1045.3854	1045.3842	1.18	$[M2-H]^{2-}$	NeuAcHexNAc3Hex-core2AB
12.76	1016.8737	1016.8734	0.26	$[M2-H]^{2-}$	dHexNeuAcHexNAc2Hex-core2AB
12.96	996.8683	996.8683	2.06	$[M2-H]^{2-}$	dHex3HexNAcHex2-core2AB
13.14	1025.3787	1025.3811	2.33	$[M2-H]^{2-}$	dHex2HexNAc2Hex2-core2AB
13.92	1024.8705	1024.8709	0.38	$[M2-H]^{2-}$	NeuAcHexNAc2Hex2-core2AB
13.81	1097.8998	1097.8998	0.04	$[M2-H]^{2-}$	dHexNeuAcHexNAc2Hex2-core2AB
14.37	1004.8657	1004.8678	2.1	$[M2-H]^{2-}$	dHex2HexNAcHex3-core2AB
14.70	1134.9176	1134.9182	0.56	$[M2-H]^{2-}$	dHexHexNAc3Hex3-core2AB
14.46	1098.4093	1098.4100	0.7	$[M2-H]^{2-}$	dHex3HexNAc2Hex2-core2AB
15.81	1170.9277	1170.9288	0.94	$[M2-H]^{2-}$	dHex2NeuAcHexNAc2Hex2-core2AB
15.66	1077.8960	1077.8968	0.71	$[M2-H]^{2-}$	dHex3HexNAcHex3-core2AB
15.51	1171.4375	1171.439	1.28	$[M2-H]^{2-}$	dHex4HexNAc2Hex2-core2AB
15.70	1207.9463	1207.9472	0.74	$[M2-H]^{2-}$	dHex2HexNAc3Hex3-core2AB
16.07	1199.4423	1199.4395	2.3	$[M2-H]^{2-}$	dHexNeuAcHexNAc3Hex2-core2AB
16.46	1077.3869	1077.3866	0.3	$[M2-H]^{2-}$	dHexNeuAcHexNAcHex2-core2AB
16.65	1208.4482	1208.4574	7.6	$[M2-H]^{2-}$	dHexHexNAc2MeHex4Pen-core2AB
16.72	1171.4375	1171.439	1.28	$[M2-H]^{2-}$	dHex4HexNAc2Hex2-core2AB
17.03	1207.9463	1207.9472	0.74	$[M2-H]^{2-}$	dHex2HexNAc3Hex3-core2AB
17.03	1150.4171	1150.4155	1.37	$[M2-H]^{2-}$	dHex2NeuAcHexNAcHex3-core2AB
17.01	1272.9785	1272.9787	0.14	$[M2-H]^{2-}$	dHex4HexNAc3Hex2-core2AB
17.30	1000.3504	1000.3495	0.94	$[M2-H]^{2-}$	Hex6-core2AB
17.38	1170.9277	1170.9288	0.94	$[M2-H]^{2-}$	dHex2NeuAcHexNAc2Hex2-core2AB

17.30	1280.4664	1280.4659	0.35	$[M2-H]^{2-}$	dHexNeuAcHexNAc3Hex3-core2AB
17.90	1280.974	1280.9761	1.67	$[M2-H]^{2-}$	dHex3HexNAc3Hex3-core2AB
18.57	1243.9556	1243.9578	1.73	$[M2-H]^{2-}$	dHex3NeuAcHexNAc2Hex2-core2AB
18.72	1345.4998	1345.4974	1.75	$[M2-H]^{2-}$	dHex3NeuAcHexNAc3Hex2-core2AB
18.76	1353.4962	1353.4949	0.07	$[M2-H]^{2-}$	dHex2NeuAcHexNAc3Hex3-core2AB
19.04	1243.4486	1243.4476	0.84	$[M2-H]^{2-}$	dHexNeuAc2HexNAc2Hex2-core2AB
19.11	1316.9874	1316.9867	0.52	$[M2-H]^{2-}$	dHex4NeuAcHexNAc2Hex2-core2AB
19.70	1354.0061	1354.0051	0.73	$[M2-H]^{2-}$	dHex4HexNAc3Hex3-core2AB
19.72	1344.9885	1344.9872	0.94	$[M2-H]^{2-}$	dHexNeuAc2HexNAc3Hex2-core2AB
19.74	1426.5237	1426.5239	0.1	$[M2-H]^{2-}$	dHex3NeuAcHexNAc3Hex3-core2AB
19.95	1316.4779	1316.4765	1.05	$[M2-H]^{2-}$	dHex2NeuAc2HexNAc2Hex2-core2AB
19.97	1345.4998	1345.4974	1.75	$[M2-H]^{2-}$	dHex3NeuAcHexNAc3Hex2-core2AB
20.10	1353.4960	1353.4949	0.81	$[M2-H]^{2-}$	dHex2NeuAcHexNAc3Hex3-core2AB
20.49	1389.5063	1389.5055	0.6	$[M2-H]^{2-}$	dHex3NeuAc2HexNAc2Hex2-core2AB
20.63	1354.0039	1354.0051	0.88	$[M2-H]^{2-}$	dHex4HexNAc3Hex3-core2AB
20.63	1426.5252	1426.5239	0.94	$[M2-H]^{2-}$	dHex3NeuAcHexNAc3Hex3-core2AB
20.91	1426.5237	1426.5239	0.1	$[M2-H]^{2-}$	dHex3NeuAcHexNAc3Hex3-core2AB
21.24	1499.0454	1499.0426	1.83	$[M2-H]^{2-}$	dHex2NeuAc2HexNAc3Hex3-core2AB
21.99	1500.0641	1500.0504	9.11	$[M2-H]^{2-}$	dHexHexNAc5Hex5-core2AB

22.10	1426.5237	1426.5239	0.1	$[M2-H]^{2-}$	dHex3NeuAcHexNAc3Hex3-core2AB
22.16	1499.5560	1499.5528	2.06	$[M2-H]^{2-}$	dHex3NeuAcHexNAc2Hex2-core2AB
22.41	1427.0353	1427.0341	0.87	$[M2-H]^{2-}$	HexNAc5Hex5-core2AB
22.82	1528.0257	1528.0635	4.7	$[M2-H]^{2-}$	dHex3NeuAcHexNAc4Hex3-core2AB
23.39	1572.5830	1572.5692	0.6	$[M2-H]^{2-}$	dHex2NeuAc3HexNAc3Hex2-core2AB
23.83	1573.0929	1573.0794	8.59	$[M2-H]^{2-}$	dHex2HexNAc5Hex5-core2AB
	1048.3952	1048.3939	0.84	$[M2-H]^{2-}$	dHexNeuAcHexNAc3-core2AB
	1573.0920	1573.0794	1.1	$[M2-H]^{2-}$	dHex2HexNAc5Hex5-core2AB
23.85	1596.0695	1596.0639	1.5	$[M2-H]^{2-}$	dHexNeuAc2HexNAc2Hex3-core2AB
23.87	1601.0591	1601.0869	1.2	$[M2-H]^{2-}$	dHexNeuAc3HexNAc4Hex2-core2AB
24.17	1572.0744	1572.0716	1.8	$[M2-H]^{2-}$	dHex4NeuAc2HexNAc3Hex2-core2AB
24.45	1645.1060	1645.1005	3.33	$[M2-H]^{2-}$	dHex4NeuAc2HexNAc3Hex3-core2AB
	1645.6101	1645.5981	7.27	$[M2-H]^{2-}$	dHexNeuAcHexNAc5Hex5-core2AB
26.26	1145.0822	1145.0839	1.47	$[M3-H]^{3-}$	dHex5NeuAc2HexNAc3Hex3
26.33	1144.7455	1144.7438	1.52	$[M3-H]^{3-}$	dHex3NeuAc3HexNAc3Hex3
26.40	1193.4317	1193.4297	1.65	$[M3-H]^{3-}$	dHex4NeuAc3HexNAc3Hex3-core2AB
27.84	1218.4469	1218.4488	1.54	$[M3-H]^{3-}$	dHex3NeuAcHeNAc4Hex4
	1218.4469	1218.4404	5.33	$[M3-H]^{3-}$	dHexNeuNAcHexNAc6Hex6

Table 16. List of proteins quantified by SWATH analysis of saliva from healthy controls and patients with PVL, early and advanced OSCC stages. The UniProt protein ID, the protein name, and subcellular location are presented. The representative identification with the number (#) of peptides with which a protein was determined and quantification data (log2Med)) are given in each case. EPR denotes endoplasmic reticulum, MIM- mitochondrial inner membrane, MOM- mitochondrial outer membrane.

Index	Protein ID	Protein name	Ctrl	PVL	eOSCC	advOSCC	# peptides	Main subcellular location
1	P61604	10 kDa heat shock protein	-0,139	-0,202	-0,186	-0,758	2	Mitochondria
2	P31946	14-3-3 protein beta/alpha	0,114	-0,261	0,178	0,071	2	Cytoplasm
3	P62258	14-3-3 protein epsilon	1,449	1,522	1,405	1,579	5	Nucleus
4	P61981	14-3-3 protein gamma	-0,239	-0,198	-0,397	0,101	4	Cytoplasm
5	P31947	14-3-3 protein sigma	3,893	4,303	4,284	4,253	12	Cytoplasm
6	P27348	14-3-3 protein theta	-0,907	-0,443	-0,703	-0,592	3	Cytoplasm
7	P63104	14-3-3 protein zeta/delta	4,632	4,677	4,532	4,668	10	Cytoplasm
8	Q99943	1-acyl-sn-glycerol-3-phosphate acyltransferase alpha	3,033	1,998	2,656	0,753	1	EPR membrane
9	Q99460	26S proteasome non-ATPase regulatory subunit 1	-3,307	-2,457	-2,899	-3,092	1	-
10	O43242	26S proteasome non-ATPase regulatory subunit 3	-4,128	-4,682	-4,310	-4,528	1	-
11	Q15008	26S proteasome non-ATPase regulatory subunit 6	-2,840	-1,872	-2,488	-2,661	1	-
12	P51665	26S proteasome non-ATPase regulatory subunit 7	-3,304	-3,174	-3,240	-3,053	1	-
13	Q99714	3-hydroxyacyl-CoA dehydrogenase type-2	-0,120	-0,260	-0,104	-0,407	2	Mitochondria
14	Q6NVY1	3-hydroxyisobutyryl-CoA hydrolase	-3,720	-3,175	-3,332	-3,467	1	Mitochondria
15	P46783	40S ribosomal protein S10	0,081	0,111	-0,138	-0,406	3	Cytoplasm
16	P25398	40S ribosomal protein S12	-2,261	-2,463	-1,965	-2,216	1	Cytoplasm
17	P62277	40S ribosomal protein S13	-0,898	-1,348	-1,599	-1,935	2	-
18	P62263	40S ribosomal protein S14	-2,109	-2,208	-2,337	-2,743	1	-
19	P62244	40S ribosomal protein S15a	0,668	0,844	0,606	0,151	3	-
20	P62249	40S ribosomal protein S16	1,113	1,344	1,043	0,921	5	-
21	P08708	40S ribosomal protein S17	-2,180	-1,844	-2,006	-2,355	2	-
22	P62269	40S ribosomal protein S18	1,299	1,391	1,247	0,749	5	Cytoplasm
23	P15880	40S ribosomal protein S2	0,948	0,977	0,714	0,953	5	-
24	P60866	40S ribosomal protein S20	0,610	0,741	0,625	0,288	2	Cytoplasm
25	P62266	40S ribosomal protein S23	-0,382	-0,253	-0,652	-0,756	2	Cytoplasm
26	P62851	40S ribosomal protein S25	0,739	0,841	0,600	0,238	2	-
27	P62854	40S ribosomal protein S26	-1,926	-1,930	-1,977	-2,282	2	Cytoplasm
28	P62857	40S ribosomal protein S28	3,068	0,291	-0,902	-1,023	1	Cytoplasm
29	P62273	40S ribosomal protein S29	-1,487	-1,830	-1,918	-2,524	1	Cytoplasm
30	P23396	40S ribosomal protein S3	1,537	1,853	1,398	1,557	5	Cytoplasm
31	P61247	40S ribosomal protein S3a	0,805	1,020	0,914	1,018	4	Cytoplasm
32	P62701	40S ribosomal protein S4	0,987	1,513	1,187	1,063	6	Cytoplasm
33	P46782	40S ribosomal protein S5	-0,823	-0,413	-0,516	-0,129	2	-
34	P62753	40S ribosomal protein S6	-0,450	-0,083	-0,184	-0,332	2	-
35	P62081	40S ribosomal protein S7	-0,002	0,205	0,016	-0,301	2	Cytoplasm

36	P62241	40S ribosomal protein S8	-0,060	0,551	0,159	0,349	4	Cytoplasm
37	P46781	40S ribosomal protein S9	0,033	0,411	0,378	0,238	2	Cytoplasm
38	P08865	40S ribosomal protein SA	0,854	1,006	0,546	0,607	4	Cell membrane
39	P49189	4-trimethylaminobutyraldehyde dehydrogenase	-0,029	-0,407	-0,222	-0,136	3	Cytoplasm
40	P10809	60 kDa heat shock protein, mitochondrial	-2,011	-2,120	-1,969	-1,783	3	Mitochondria
41	P05388	60S acidic ribosomal protein P0	0,997	1,039	0,987	0,899	5	Nucleus
42	P05387	60S acidic ribosomal protein P2	-0,688	-0,503	-0,859	-0,968	2	-
43	P62913	60S ribosomal protein L11	-0,197	0,042	-0,253	-0,463	1	Nucleus
44	P30050	60S ribosomal protein L12	0,146	0,094	0,174	-0,170	2	-
45	P26373	60S ribosomal protein L13	-0,098	-0,095	-0,566	-0,678	3	-
46	P40429	60S ribosomal protein L13a	-1,701	-1,343	-1,436	-1,286	1	Cytoplasm
47	P50914	60S ribosomal protein L14	-0,235	0,028	-0,065	0,047	2	-
48	P61313	60S ribosomal protein L15	-2,346	-1,778	-2,052	-2,922	1	Membrane
49	P18621	60S ribosomal protein L17	-2,438	-2,797	-2,750	-3,031	1	-
50	Q07020	60S ribosomal protein L18	0,621	0,561	0,630	0,388	3	Cytoplasm
51	Q02543	60S ribosomal protein L18a	-0,917	-0,504	-0,719	-1,294	2	-
52	P46778	60S ribosomal protein L21	-2,985	-2,504	-3,752	-2,553	1	Cytoplasm
53	P35268	60S ribosomal protein L22	0,409	0,286	0,399	-0,058	2	-
54	P62829	60S ribosomal protein L23	0,860	0,802	0,794	0,609	3	-
55	P61254	60S ribosomal protein L26	-1,435	-0,662	-0,937	-1,270	2	-
56	P61353	60S ribosomal protein L27	0,302	0,207	0,016	-0,103	1	Cytoplasm
57	P46776	60S ribosomal protein L27a	-0,052	0,041	-0,607	-0,539	1	-
58	P46779	60S ribosomal protein L28	-0,251	0,029	-0,280	-0,632	2	-
59	P39023	60S ribosomal protein L3	-2,555	-2,842	-2,889	-3,409	1	Nucleus
60	P62888	60S ribosomal protein L30	0,569	0,699	0,386	0,442	3	-
61	P49207	60S ribosomal protein L34	0,094	0,147	0,194	-0,073	2	Cytoplasm
62	Q969Q0	60S ribosomal protein L36a	-2,128	-2,077	-2,351	-2,701	1	Cytoplasm
63	P36578	60S ribosomal protein L4	0,337	0,701	0,119	0,478	3	-
64	P46777	60S ribosomal protein L5	-1,527	-0,940	-1,373	-1,496	1	Cytoplasm
65	Q02878	60S ribosomal protein L6	-0,849	-1,260	-0,889	-2,008	1	Cytoplasm
66	P18124	60S ribosomal protein L7	0,892	0,669	0,882	0,034	2	-
67	P62424	60S ribosomal protein L7a	0,019	-0,431	-0,272	0,126	3	-
68	P52209	6-phosphogluconate dehydrogenase	4,441	4,493	4,215	4,212	18	Cytoplasm
69	O95336	6-phosphogluconolactonase	-0,490	-1,498	-0,775	-0,923	3	Cytoplasm
70	Q13510	Acid ceramidase	-1,853	-2,639	-2,085	-2,144	1	Lysosome
71	P68032	Actin, alpha cardiac muscle 1	4,151	3,573	3,826	3,920	4	Cytoplasm
72	P61160	Actin-related protein 2	0,009	0,182	-0,003	-0,532	2	Cytoplasm
73	Q92747	Actin-related protein 2/3 complex subunit 1A	-2,783	-2,917	-2,612	-2,980	1	Cytoplasm
74	O15143	Actin-related protein 2/3 complex subunit 1B	0,651	0,706	0,251	0,155	2	Cytoplasm
75	O15144	Actin-related protein 2/3 complex subunit 2	0,132	0,366	0,262	0,021	3	Cytoplasm
76	O15145	Actin-related protein 2/3 complex subunit 3	-1,895	-1,599	-3,141	-1,327	1	Cytoplasm

77	P59998	Actin-related protein 2/3 complex subunit 4	2,891	2,911	2,688	2,725	6	Cytoplasm
78	O15511	Actin-related protein 2/3 complex subunit 5	0,857	0,623	0,356	0,175	1	Cytoplasm
79	P61158	Actin-related protein 3	-0,045	0,082	-0,062	0,010	4	Cytoplasm
80	P07108	Acyl-CoA-binding protein	-2,703	-2,055	-2,166	-2,448	2	EPR
81	O75608	Acyl-protein thioesterase 1	-2,886	-2,423	-2,237	-2,742	1	Cytoplasm
82	P23526	Adenosylhomocysteinase	-1,156	-1,389	-1,473	-1,364	2	Cytoplasm
83	Q01518	Adenylyl cyclase-associated protein 1	1,871	1,790	1,907	2,126	13	Cell membrane
84	Q9HDC9	Adipocyte plasma membrane-associated protein	-1,779	-0,965	-1,281	-1,815	1	Membrane
85	P05141	ADP/ATP translocase 2	1,065	1,396	0,922	1,165	2	MIM
86	P12236	ADP/ATP translocase 3	-1,880	-1,265	-1,916	-1,439	1	MIM
87	P61204	ADP-ribosylation factor 3	-1,423	-0,901	-0,882	-1,694	1	Golgi apparatus
88	P62330	ADP-ribosylation factor 6	0,244	0,995	0,865	0,507	3	Cytoplasm
89	Q9NVJ2	ADP-ribosylation factor-like protein 8B	-3,018	-2,482	-2,372	-3,259	1	Membrane
90	P30838	Aldehyde dehydrogenase	2,520	3,145	2,586	2,587	8	Cytoplasm
91	P05091	Aldehyde dehydrogenase	-2,340	-2,084	-2,865	-2,609	1	Mitochondria
92	P51648	Aldehyde dehydrogenase family 3 member A2	-1,707	-0,978	-1,511	-1,937	1	Microsome membrane
93	P14550	Aldo-keto reductase family 1 member A1	-0,167	-0,329	-0,347	0,014	3	Cytoplasm
94	O60218	Aldo-keto reductase family 1 member B10	0,810	1,599	1,290	1,555	7	Lysosome
95	Q04828	Aldo-keto reductase family 1 member C1	-0,762	0,419	-0,557	-0,073	1	Cytoplasm
96	P40394	All-trans-retinol dehydrogenase [NAD	2,005	2,323	1,873	1,701	9	Cytoplasm
97	P02763	Alpha-1-acid glycoprotein 1	1,831	2,414	3,008	4,106	3	Secreted
98	P19652	Alpha-1-acid glycoprotein 2	-0,258	0,121	0,947	1,289	2	Secreted
99	P01011	Alpha-1-antichymotrypsin	1,740	1,908	2,214	3,117	9	Secreted
100	P01009	Alpha-1-antitrypsin	4,865	5,399	5,809	6,903	27	Secreted
101	P04217	Alpha-1B-glycoprotein	1,052	1,122	1,845	2,637	7	Secreted
102	P08697	Alpha-2-antiplasmin	-3,477	-3,849	-3,292	-2,839	1	Secreted
103	P02765	Alpha-2-HS-glycoprotein	0,474	0,993	1,639	2,101	3	Secreted
104	P01023	Alpha-2-macroglobulin	4,077	4,419	5,188	6,078	42	Secreted
105	A8K2U0	Alpha-2-macroglobulin-like protein 1	5,161	5,073	4,778	5,083	30	Secreted
106	P12814	Alpha-actinin-1	-0,107	-0,122	-0,336	-0,135	5	Cytoplasm
107	Q08043	Alpha-actinin-3	-2,419	-2,839	-2,988	-3,257	2	Cytoplasm
108	O43707	Alpha-actinin-4	2,274	2,253	2,213	2,285	14	Nucleus
109	P04745	Alpha-amylase 1	9,523	8,984	8,813	9,360	19	Secreted
110	P61163	Alpha-centractin	-2,657	-2,625	-2,374	-2,751	2	Cytoplasm
111	P02511	Alpha-crystallin B chain	0,974	0,906	0,699	0,536	6	Cytoplasm
112	P06733	Alpha-enolase	4,535	4,392	4,508	4,579	27	Cytoplasm
113	P54920	Alpha-soluble NSF attachment protein	-2,175	-1,231	-2,503	-1,798	2	Cell membrane
114	Q13155	Aminoacyl tRNA synthase protein 2	-4,160	-3,452	-3,472	-2,996	1	Cytoplasm
115	Q9UBD6	Ammonium transporter Rh type C	1,409	1,827	0,790	0,841	4	Cell membrane
116	P01019	Angiotensinogen	-1,473	-1,190	-0,544	0,195	4	Secreted

117	P04083	Annexin A1	7,863	8,321	7,768	7,462	28	Nucleus
118	P50995	Annexin A11	0,219	0,762	0,247	-0,210	4	Cytoplasm
119	P07355	Annexin A2	2,979	3,369	2,944	2,797	6	Secreted
120	P12429	Annexin A3	4,741	4,755	4,302	4,316	16	Cytoplasm
121	P09525	Annexin A4	0,593	1,138	0,891	0,580	5	Cell surface
122	P08758	Annexin A5	1,899	2,077	1,935	1,880	6	Cytoplasm
123	P08133	Annexin A6	2,235	2,376	2,016	1,971	9	Cytoplasm
124	P20073	Annexin A7	-2,439	-2,286	-2,152	-2,778	1	Cytoplasm
125	Q5VT79	Annexin A8-like protein 1	2,022	2,433	2,117	2,017	7	-
126	P03973	Antileukoproteinase	0,475	0,607	1,038	0,488	4	Secreted
127	P01008	Antithrombin-III	0,863	1,204	1,424	2,292	6	Secreted
128	P02647	Apolipoprotein A-I	3,208	3,908	5,127	5,472	23	Secreted
129	P02652	Apolipoprotein A-II	-2,835	-2,939	-1,572	-1,937	1	Secreted
130	P06727	Apolipoprotein A-IV	-1,710	-2,472	-1,502	-1,363	2	Secreted
131	P04114	Apolipoprotein B-100	1,680	1,728	1,957	2,658	12	Cytoplasm
132	P05090	Apolipoprotein D	-1,578	-0,754	-0,700	-0,466	3	Secreted
133	Q9ULZ3	Apoptosis-associated speck-like protein	-3,518	-3,858	-3,910	-3,614	1	Cytoplasm
134	P18054	Arachidonate 12-lipoxygenase, 12S-type	-3,083	-2,321	-2,983	-2,993	2	Cytoplasm
135	P09917	Arachidonate 5-lipoxygenase	-3,396	-2,922	-3,039	-2,647	1	Nucleus membrane
136	P20292	Arachidonate 5-lipoxygenase-activating protein	0,605	0,690	-0,626	-0,544	1	Cytoplasm
137	P05089	Arginase-1	-1,897	-1,972	-2,303	-2,476	2	Mitochondria
138	P00505	Aspartate aminotransferase	-4,517	-4,787	-5,508	-4,780	1	Cytoplasm
139	P17174	Aspartate aminotransferas,	-0,916	-0,959	-0,600	-0,427	3	Cytoplasm
140	P24539	ATP synthase F	-4,774	-5,257	-4,973	-4,528	1	Mitochondria
141	P25705	ATP synthase subunit alpha	3,488	3,272	3,609	3,314	12	MIM
142	P06576	ATP synthase subunit beta	2,673	2,968	2,751	2,708	12	Mitochondria
143	O75964	ATP synthase subunit g, mitochondrial	-0,408	-0,345	-0,147	-0,769	3	MIM
144	P36542	ATP synthase subunit gamma	-2,380	-1,852	-1,591	-1,694	2	Mitochondria
145	P48047	ATP synthase subunit O	-0,423	-0,857	-0,484	-0,446	2	Mitochondria
146	P17858	ATP-dependent 6-phosphofructokinase	-4,313	-4,651	-4,339	-4,356	1	Cell membrane
147	O00571	ATP-dependent RNA helicase	-1,962	-1,402	-2,081	-1,917	3	Cytoplasm
148	P20160	Azurocidin	5,354	5,491	4,805	5,156	11	Secreted
149	P17213	Bactericidal permeability-increasing protein	0,449	0,408	0,452	0,414	4	Secreted
150	P04280	Basic salivary proline-rich protein 1	-9,246	-10,681	-8,802	-9,144	1	Secreted
151	P02812	Basic salivary proline-rich protein 2	-4,800	-7,951	-7,139	-5,815	1	Secreted
152	Q04118	Basic salivary proline-rich protein 3	-1,161	-4,562	-3,513	-3,056	2	Secreted
153	P10163	Basic salivary proline-rich protein 4	-1,583	-2,436	-2,271	-1,946	3	EPR membrane
154	P51572	B-cell receptor-associated protein 31	-0,987	-0,958	-0,877	-0,765	4	Secreted
155	P02749	Beta-2-glycoprotein 1	0,505	0,325	1,151	2,174	6	Secreted
156	P61769	Beta-2-microglobulin	1,466	1,615	1,785	1,609	5	Cytoplasm

157	P53004	Biliverdin reductase A	-2,202	-1,377	-1,421	-2,155	2	Secreted
158	Q9NP55	BPI fold-containing family A member 1	2,607	3,679	3,335	2,508	8	Secreted
159	Q96DR5	BPI fold-containing family A member 2	6,460	5,640	6,031	5,331	21	Secreted
160	Q8TDL5	BPI fold-containing family B member 1	4,896	5,052	4,703	4,625	24	Secreted
161	Q8N4F0	BPI fold-containing family B member 2	4,650	4,393	4,445	4,973	11	Cytoplasm
162	Q5TH69	Brefeldin A-inhibited guanine nucleotide-exchange protein 3	1,023	0,575	0,490	-0,178	1	Secreted
163	P04003	C4b-binding protein alpha chain	-2,122	-1,934	-1,664	-0,772	3	Cell junction
164	P12830	Cadherin-1	-0,334	-0,644	-0,995	-0,807	2	EPR membrane
165	Q9UM00	Calcium load-activated calcium channel	-2,749	-2,926	-2,836	-3,399	1	Cell membrane
166	Q14CN2	Calcium-activated chloride channel regulator 4	0,696	0,725	0,802	0,927	6	Cytoplasm
167	Q9Y2V2	Calcium-regulated heat-stable protein 1	-3,059	-3,130	-2,306	-3,308	1	Cytoplasm
168	P0DP25	Calmodulin-3	-2,374	-1,632	-1,632	-2,855	1	
169	P27482	Calmodulin-like protein 3	0,673	1,562	1,215	0,750	3	
170	Q9NZT1	Calmodulin-like protein 5	-2,158	-0,377	-0,866	-3,162	1	EPR membrane
171	P27824	Calnexin	0,361	0,500	0,591	0,135	4	Cytoplasm
172	P04632	Calpain small subunit 1	1,057	1,545	1,415	1,060	4	Cytoplasm
173	P07384	Calpain-1 catalytic subunit	2,541	2,877	2,468	2,640	11	Cytoplasm; Cell membrane
174	P17655	Calpain-2 catalytic subunit	-0,648	-0,312	-0,414	-0,624	3	EPR lumen
175	P27797	Calreticulin	1,824	2,093	2,006	1,948	11	Cytoplasm
176	P00915	Carbonic anhydrase 1	-0,164	-0,117	-0,375	1,578	5	Secreted
177	P23280	Carbonic anhydrase 6	6,560	5,485	5,657	5,815	18	Cell membrane
178	P40199	Carcinoembryonic antigen-related cell adhesion molecule 6	-0,508	-0,295	-0,561	-0,363	2	Cell membrane
179	P31997	Carcinoembryonic antigen-related cell adhesion molecule 8	-3,968	-4,191	-5,370	-3,884	1	Cytoplasm
180	P31944	Caspase-14	-2,916	-1,439	-1,225	-2,766	2	Peroxisome
181	P04040	Catalase	2,138	1,487	1,407	1,798	10	Secreted
182	P49913	Cathelicidin antimicrobial peptide	3,276	3,170	2,744	2,890	5	Lysosome
183	P07858	Cathepsin B	0,553	-0,352	0,116	0,189	2	Secreted
184	P07339	Cathepsin D	1,841	1,515	2,279	1,410	3	Cell surface
185	P08311	Cathepsin G	6,028	6,138	5,407	5,491	13	Secreted
186	Q9NRJ3	C-C motif chemokine 28	-0,740	-1,559	-1,274	-1,664	1	Cell membrane
187	P13987	CD59 glycoprotein	-1,469	-1,406	-1,338	-1,583	2	Cell membrane
188	P21926	CD9 antigen	-0,282	-0,261	-0,868	-0,900	2	Cell membrane
189	P60953	Cell division control protein 42 homolog	1,257	0,990	0,796	1,151	5	Cytoplasm
190	P29373	Cellular retinoic acid-binding protein 2	0,012	0,193	-0,066	0,227	3	Secreted
191	P00450	Ceruloplasmin	1,795	2,058	2,567	3,111	13	Secreted

192	Q15782	Chitinase-3-like protein 2	-1,683	-1,519	-2,157	-0,729	2	Nucleus
193	O00299	Chloride intracellular channel protein 1	1,711	1,958	1,820	1,894	8	Cell membrane
194	O95833	Chloride intracellular channel protein 3	-0,545	-0,530	-0,473	-0,538	3	Mitochondria
195	O75390	Citrate synthase	1,449	1,236	1,420	1,481	5	Mitochondria
196	Q00610	Clathrin heavy chain 1	0,545	1,325	0,621	0,768	7	
197	P10909	Clusterin	1,670	1,190	1,761	2,297	5	Cytoplasm
198	Q14019	Coactosin-like protein	-0,399	-1,201	-0,961	-0,613	2	Cytoplasm
199	Q9Y678	Coatomer subunit gamma-1	-3,691	-2,978	-3,586	-3,265	1	EPR lumen
200	O00748	Cocaine esterase	-2,580	-2,920	-2,640	-2,677	1	Nucleus matrix
201	P23528	Cofilin-1	1,988	1,648	1,871	1,843	6	Secreted
202	P02747	Complement C1q subcomponent subunit C	-3,678	-2,894	-2,522	-1,987	1	Secreted
203	Q9BXJ4	Complement C1q tumour necrosis factor-related protein 3	-0,066	-0,809	-0,904	-0,744	3	Secreted
204	P01024	Complement C3	4,757	5,217	5,779	6,463	67	Secreted
205	P0C0L4	Complement C4-A	2,527	2,496	3,250	3,847	23	Secreted
206	P01031	Complement C5	-3,257	-3,877	-2,817	-1,830	1	Secreted
207	P00751	Complement factor B	0,808	1,256	1,738	2,599	11	Secreted
208	P08603	Complement factor H	0,668	0,360	0,836	1,352	7	Secreted
209	P05156	Complement factor I	-3,370	-3,252	-3,014	-2,457	2	Nucleus
210	O75131	Copine-3	0,160	0,506	-0,047	0,124	3	Nucleus
211	O75367	Core histone macro-H2A.1	0,521	0,635	0,355	0,343	2	Cytoplasm
212	Q9BYD5	Cornifelin	-0,399	0,218	-0,536	-0,660	2	Cytoplasm
213	Q9UBG3	Cornulin	4,408	4,438	4,172	4,178	23	Cytoplasm
214	P31146	Coronin-1A	2,814	1,895	1,860	2,085	9	Cytoplasm
215	P06732	Creatine kinase M-type	-4,112	-4,508	-4,366	-4,487	1	MIM
216	P12532	Creatine kinase U-type	-0,349	-0,903	-0,969	-1,193	2	Mitochondria
217	P01040	Cystatin-A	2,094	2,795	2,179	1,187	7	Cytoplasm
218	P04080	Cystatin-B	4,799	5,116	4,731	4,414	11	Cytoplasm
219	P01034	Cystatin-C	3,147	1,569	1,930	2,007	9	Secreted
220	P28325	Cystatin-D	4,669	4,489	4,422	4,326	14	Secreted
221	P01036	Cystatin-S	3,679	1,155	0,485	1,460	2	Secreted
222	P09228	Cystatin-SA	4,382	3,735	3,447	3,445	8	Secreted
223	P01037	Cystatin-SN	7,015	5,803	5,836	5,540	7	Secreted
224	P54108	Cysteine-rich secretory protein 3	4,332	3,472	3,700	3,696	9	Secreted
225	P04839	Cytochrome b-245 heavy chain	0,909	1,021	0,515	0,129	4	Cell membrane
226	P31930	Cytochrome b-c1 complex subunit 1	-1,705	-1,662	-1,555	-1,844	2	MIM
227	P22695	Cytochrome b-c1 complex subunit 2	-4,117	-4,225	-3,540	-4,388	1	MIM
228	P99999	Cytochrome c	-4,375	-4,150	-3,451	-3,565	1	Mitochondria
229	P00403	Cytochrome c oxidase subunit 2	-0,106	0,442	0,307	-0,407	1	MIM
230	P13073	Cytochrome c oxidase subunit 4 isoform 1	-1,016	-0,822	-1,206	-1,232	1	MIM

231	Q7L576	Cytoplasmic FMR1-interacting protein 1	-3,111	-3,037	-2,819	-2,650	1	Cytoplasm
232	Q07065	Cytoskeleton-associated protein 4	-1,490	-1,629	-1,401	-1,434	3	EPR membrane
233	P49902	Cytosolic purine 5'-nucleotidase	-3,298	-3,437	-3,824	-3,482	1	Cytoplasm
234	Q96LJ7	Dehydrogenase/reductase SDR family member 1	-1,860	-1,217	-2,000	-2,038	1	
235	Q9UGM3	Deleted in malignant brain tumours 1 protein	5,273	5,654	5,121	4,959	19	Secreted
236	Q14739	Delta	-2,337	-2,616	-2,430	-3,140	1	Nucleus
237	P81605	Dermcidin	-3,029	-2,935	-1,424	-3,641	1	Secreted
238	Q02487	Desmocollin-2	1,419	1,699	1,250	1,044	5	Cell membrane
239	Q02413	Desmoglein-1	-1,221	-1,375	-1,047	-1,439	2	Cell membrane
240	P32926	Desmoglein-3	1,779	1,507	1,334	1,446	11	Cell membrane
241	P15924	Desmoplakin	3,312	4,544	3,705	3,340	29	Cell junction
242	P60981	Destrin	-1,620	-1,297	-1,678	-1,587	1	
243	P36957	Dihydrolipoyllysine-residue succinyltransferase	2,378	0,307	2,470	-0,944	1	Mitochondria
244	P53634	Dipeptidyl peptidase 1	-2,178	-1,909	-2,019	-2,907	1	Lysosome
245	P25685	DnaJ homolog subfamily B member 1	-2,459	-2,416	-1,872	-2,069	1	Cytoplasm
246	P04843	Dolichyl-diphosphooligosaccharide subunit 1	-2,154	-1,761	-1,857	-2,442	2	EPR
247	P04844	Dolichyl-diphosphooligosaccharide subunit 2	-1,527	-1,412	-1,411	-2,039	2	EPR
248	O95147	Dual specificity protein phosphatase 14	-4,980	-5,281	-4,787	-5,043	1	
249	P63167	Dynein light chain 1	0,875	1,152	1,058	0,885	3	Cytoplasm
250	P24534	Elongation factor 1-beta	-3,346	-3,647	-2,863	-3,300	1	
251	P29692	Elongation factor 1-delta	-1,925	-1,335	-1,527	-0,678	2	
252	P26641	Elongation factor 1-gamma	0,805	1,048	0,850	0,749	4	
253	P13639	Elongation factor 2	2,888	3,325	3,001	3,142	14	Cytoplasm
254	P49411	Elongation factor Tu	-1,786	-3,586	-2,064	-3,465	1	Mitochondria
255	P11021	Endoplasmic reticulum chaperone BiP	2,160	1,886	2,239	2,409	14	EPR lumen
256	P30040	Endoplasmic reticulum resident protein 29	-2,530	-2,404	-1,815	-2,748	1	EPR lumen
257	P14625	Endoplasmin	-0,208	-0,076	0,018	0,320	3	EPR lumen
258	Q92817	Envoplakin	-0,914	-1,066	-1,117	-0,676	3	Cell junction
259	P12724	Eosinophil cationic protein	2,428	2,131	2,261	2,172	5	Secreted
260	P11678	Eosinophil peroxidase	-3,355	-3,768	-3,166	-2,745	2	Cytoplasm
261	Q8N3Y7	Epidermal retinol dehydrogenase 2	-0,103	0,481	-0,213	0,038	2	EPR membrane
262	P07099	Epoxide hydrolase 1	-0,768	-0,104	-0,338	-0,494	3	Microsome membrane
263	Q96HE7	ERO1-like protein alpha	3,040	3,265	3,136	3,240	14	EPR membrane
264	P27105	Erythrocyte band 7 integral membrane protein	0,652	0,516	0,212	0,584	3	Cell membrane
265	P60842	Eukaryotic initiation factor 4A-I	1,396	1,765	1,759	1,657	5	

266	P62495	Eukaryotic peptide chain release factor subunit 1	-2,959	-2,926	-2,940	-2,950	1	Cytoplasm
267	O00303	Eukaryotic translation initiation factor 3 subunit F	-4,550	-3,324	-4,027	-3,867	1	Cytoplasm
268	P06730	Eukaryotic translation initiation factor 4E	-4,812	-4,635	-4,324	-4,745	1	Cytoplasm
269	P63241	Eukaryotic translation initiation factor 5A-1	0,570	0,738	0,773	0,506	3	Cytoplasm
270	Q16610	Extracellular matrix protein 1	1,097	1,408	1,145	1,094	7	Secreted
271	P15311	Ezrin	0,989	1,274	0,884	1,282	5	Apical cell membrane
272	P52907	F-actin-capping protein subunit alpha-1	-0,049	0,218	-0,125	-0,116	2	Cytoplasm
273	P47756	F-actin-capping protein subunit beta	0,728	0,887	0,787	0,185	3	Cytoplasm
274	Q16658	Fascin	-1,346	-0,547	-0,855	-0,299	2	Cytoplasm
275	Q01469	Fatty acid-binding protein 5	3,207	4,273	3,925	3,542	11	Cytoplasm
276	Q6ZVX7	F-box only protein 50	0,959	1,099	1,075	1,028	5	Cytoplasm
277	P02671	Fibrinogen alpha chain	0,774	0,795	1,965	2,854	7	Secreted
278	P02675	Fibrinogen beta chain	3,037	3,288	3,995	4,671	20	Secreted
279	P02679	Fibrinogen gamma chain	3,202	3,671	4,295	4,947	16	Secreted
280	P02751	Fibronectin	-1,019	-1,114	-0,414	0,142	4	Secreted
281	P20930	Filaggrin	-0,943	0,016	-1,014	-0,738	9	Cytoplasmic granule
282	Q5D862	Filaggrin-2	-1,787	-1,036	0,233	-2,884	1	Cytoplasm
283	P21333	Filamin-A	-0,093	-0,014	-0,260	0,114	4	Cytoplasm
284	P30043	Flavin reductase	-3,603	-3,462	-3,738	-3,277	1	Cytoplasm
285	P04075	Fructose-bisphosphate aldolase A	3,326	3,323	3,372	3,534	10	Cytoplasm
286	P09972	Fructose-bisphosphate aldolase C	-1,543	-1,746	-1,637	-0,790	1	
287	P21217	Galactoside 3	-2,539	-2,534	-2,738	-2,467	1	Golgi apparatus
288	Q05315	Galectin-10	0,040	-0,419	-0,844	-0,203	3	Cytoplasm
289	Q08380	Galectin-3-binding protein	3,402	2,984	2,995	3,411	8	Secreted
290	P47929	Galectin-7	4,047	4,481	4,162	4,054	8	Cytoplasm
291	Q92820	Gamma-glutamyl hydrolase	-1,602	-1,402	-1,166	-1,250	2	Secreted
292	P17900	Ganglioside GM2 activator	-0,998	-0,731	-0,233	-1,275	2	Lysosome
293	P06396	Gelsolin	3,978	3,757	3,701	4,073	19	
294	P11413	Glucose-6-phosphate 1-dehydrogenase	2,178	2,524	2,210	2,174	8	Cytoplasm
295	P06744	Glucose-6-phosphate isomerase	3,001	2,934	2,913	2,916	10	Cytoplasm
296	P00367	Glutamate dehydrogenase 1,	-2,634	-1,453	-1,781	-1,563	1	Mitochondria
297	P15104	Glutamine synthetase	1,099	2,033	2,056	1,765	6	Cytoplasm
298	P22352	Glutathione peroxidase 3	-2,065	-2,118	-1,867	-2,006	1	Secreted
299	P78417	Glutathione S-transferase omega-1	0,301	-0,153	-0,216	0,010	1	Cytoplasm
300	P09211	Glutathione S-transferase P	4,012	4,177	4,047	4,075	9	Cytoplasm
301	P04406	Glyceraldehyde-3-phosphate dehydrogenase	4,705	4,407	4,669	4,402	18	Cytoplasm
302	P43304	Glycerol-3-phosphate dehydrogenase	-3,529	-3,704	-3,068	-2,920	1	Mitochondria
303	P13807	Glycogen synthase	-3,150	-2,115	-2,400	-2,907	1	Cytoplasm

304	P11216	Glycogen phosphorylase, brain form	-1,347	-0,781	-1,188	-2,038	1	Cytoplasm
305	P06737	Glycogen phosphorylase, liver form	0,033	-0,285	-0,595	-0,360	3	
306	Q9NZD2	Glycolipid transfer protein	-0,616	-0,149	-0,202	-0,314	3	Cytoplasm
307	P28676	Grancalcin	2,966	3,101	2,581	2,693	5	Cytoplasm
308	P01111	GTPase NRas	-1,398	-1,014	-1,754	-1,867	1	Cell membrane
309	P62826	GTP-binding nuclear protein Ran	-0,567	-0,448	-0,553	-0,301	2	Nucleus
310	Q9Y6B6	GTP-binding protein SAR1b	-1,311	-1,645	-1,395	-1,694	1	EPR membrane
311	P62879	Guanine nucleotide-binding protein G	-0,347	-0,142	-0,411	-0,863	2	Cytoplasm
312	P04899	Guanine nucleotide-binding protein G	-1,559	-1,460	-1,203	-2,005	2	Cytoplasm
313	Q6ZN66	Guanylate-binding protein 6	1,649	2,328	1,677	1,497	10	
314	P00738	Haptoglobin	3,604	3,932	4,872	6,222	25	Secreted
315	P0DMV9	Heat shock 70 kDa protein 1B	3,450	3,641	3,477	3,631	12	Cytoplasm
316	P11142	Heat shock cognate 71 kDa protein	3,009	2,921	3,080	3,200	12	Cytoplasm
317	P04792	Heat shock protein beta-1	5,214	5,235	5,152	5,055	17	Cytoplasm
318	P07900	Heat shock protein HSP 90-alpha	1,829	2,102	1,964	2,121	6	Nucleus
319	P08238	Heat shock protein HSP 90-beta	1,850	2,146	1,980	2,078	6	Cytoplasm
320	Q9Y5Z4	Heme-binding protein 2	-2,242	-2,661	-2,772	-2,559	1	Cytoplasm
321	P69905	Hemoglobin subunit alpha	2,523	2,287	2,677	5,367	10	Extracellular exosome
322	P68871	Hemoglobin subunit beta	3,433	3,610	3,837	6,542	11	Extracellular exosome
323	P02042	Hemoglobin subunit delta	-0,364	-0,736	-0,688	1,078	4	Extracellular exosome
324	P02790	Hemopexin	2,913	3,464	4,081	5,105	14	Secreted
325	P09651	Heterogeneous nuclear ribonucleoprotein A1	-1,810	-1,570	-1,272	-1,663	2	Nucleus
326	P31943	Heterogeneous nuclear ribonucleoprotein H	-1,764	-1,896	-1,588	-1,306	2	Nucleus
327	P61978	Heterogeneous nuclear ribonucleoprotein K	-0,007	0,308	0,151	-0,037	3	Cytoplasm
328	O60506	Heterogeneous nuclear ribonucleoprotein Q	-2,292	-1,529	-2,087	-2,701	1	Cytoplasm
329	P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	-2,006	-2,539	-1,807	-2,721	1	Nucleus
330	P19367	Hexokinase-1	-0,974	-0,945	-0,992	-1,231	2	MOM
331	P52790	Hexokinase-3	-1,122	-1,608	-1,320	-1,945	2	
332	P09429	High mobility group protein B1	-3,186	-3,061	-3,215	-3,097	1	Nucleus
333	P26583	High mobility group protein B2	0,088	-1,346	-1,497	-1,526	2	Nucleus
334	P15515	Histatin-1	-5,692	-5,442	-5,438	-4,585	2	Secreted
335	P04196	Histidine-rich glycoprotein	-0,295	-0,232	0,469	1,030	5	Secreted
336	P07305	Histone H1.0	-0,227	-0,773	-0,865	-1,440	3	Nucleus
337	P10412	Histone H1.4	1,337	0,293	-0,166	-0,435	3	Nucleus
338	P16401	Histone H1.5	1,755	0,936	0,671	-0,051	2	Nucleus
339	Q6FI13	Histone H2A type 2-A	-3,332	-4,403	-2,877	-3,200	1	Nucleus
340	Q8IUE6	Histone H2A type 2-B	0,215	-0,958	-0,750	-0,272	2	Nucleus
341	Q7L7L0	Histone H2A type 3	-1,029	-1,934	-1,749	-1,940	1	Nucleus

342	P06899	Histone H2B type 1-J	1,131	1,249	0,763	0,706	1	Nucleus
343	Q99877	Histone H2B type 1-N	3,878	3,933	3,454	3,354	2	Nucleus
344	P84243	Histone H3.3	-4,442	-3,897	-3,458	-4,779	1	Nucleus
345	P62805	Histone H4	8,382	8,489	7,991	7,753	11	Nucleus
346	P17693	HLA class I histocompatibility antigen	-2,268	-1,721	-1,619	-1,742	1	
347	Q86YZ3	Hornerin	-3,564	-3,392	-3,271	-3,352	3	Cytoplasm
348	P00492	Hypoxanthine-guanine phosphoribosyltransferase	-3,464	-2,772	-3,580	-3,463	1	Cytoplasm
349	Q9Y6R7	IgGFc-binding protein	3,555	2,396	2,467	2,996	21	Secreted
350	P0DOX2	Immunoglobulin alpha-2 heavy chain	5,512	4,726	4,714	5,452	11	Secreted
351	P0DOX5	Immunoglobulin gamma-1 heavy chain	3,634	3,643	4,880	4,881	8	Secreted
352	P01876	Immunoglobulin heavy constant alpha 1	8,623	7,966	8,435	8,480	16	Secreted
353	P01859	Immunoglobulin heavy constant gamma 2	4,729	5,220	5,759	6,583	6	Secreted
354	P01860	Immunoglobulin heavy constant gamma 3	0,418	0,404	1,735	2,525	4	Secreted
355	P01861	Immunoglobulin heavy constant gamma 4	-1,680	-2,553	-1,704	-0,811	1	Secreted
356	P01871	Immunoglobulin heavy constant mu	-2,670	-3,059	-2,971	-3,278	1	Secreted
357	A0A0C4DH31	Immunoglobulin heavy variable 1-18	2,150	2,107	2,370	2,393	1	Secreted
358	P01743	Immunoglobulin heavy variable 1-46	0,536	0,328	0,618	0,699	1	Secreted
359	A0A0B4J1V0	Immunoglobulin heavy variable 3-15	-0,163	0,248	-0,241	0,424	1	Secreted
360	P0DP03	Immunoglobulin heavy variable 3-30-5	0,813	0,320	0,576	0,895	1	Secreted
361	A0A0A0MS15	Immunoglobulin heavy variable 3-49	1,068	1,001	0,734	1,273	1	Secreted
362	P01780	Immunoglobulin heavy variable 3-7	1,206	1,142	1,448	1,680	1	Secreted
363	A0A0B4J1Y9	Immunoglobulin heavy variable 3-72	0,827	0,726	0,970	1,228	1	Secreted
364	A0A0B4J1X5	Immunoglobulin heavy variable 3-74	3,068	2,736	2,966	3,361	1	Secreted
365	P06331	Immunoglobulin heavy variable 4-34	2,883	2,817	3,015	2,990	2	Secreted
366	A0A0C4DH38	Immunoglobulin heavy variable 5-51	-2,799	-2,900	-2,727	-2,528	1	Secreted
367	P01591	Immunoglobulin J chain	5,981	5,513	5,813	5,642	7	Secreted
368	P0DOX7	Immunoglobulin kappa light chain	-0,952	-0,810	-0,812	-0,804	2	Secreted
369	P01594	Immunoglobulin kappa variable 1-33	-1,582	-1,783	-2,363	-1,265	1	Secreted
370	A0A0C4DH69	Immunoglobulin kappa variable 1-9	-2,134	-1,671	-1,471	-0,779	1	Secreted
371	A0A0C4DH68	Immunoglobulin kappa variable 2-24	-0,202	-0,257	-0,381	0,399	1	Secreted
372	P06310	Immunoglobulin kappa variable 2-30	-0,327	-1,188	-1,183	0,167	2	Secreted
373	P01615	Immunoglobulin kappa variable 2D-28	1,296	0,353	0,892	1,404	2	Secreted
374	P04433	Immunoglobulin kappa variable 3-11	2,989	2,372	2,787	3,327	2	Secreted
375	P01624	Immunoglobulin kappa variable 3-15	4,063	3,725	3,958	4,277	1	Secreted

376	P01619	Immunoglobulin kappa variable 3-20	1,696	1,342	1,683	2,033	1	Secreted
377	P06312	Immunoglobulin kappa variable 4-1	3,030	3,078	3,101	3,478	3	Secreted
378	A0M8Q6	Immunoglobulin lambda constant 7	-3,530	-6,338	-4,295	-5,865	1	Secreted
379	P01700	Immunoglobulin lambda variable 1-47	2,012	1,826	2,435	2,174	2	Secreted
380	P80748	Immunoglobulin lambda variable 3-21	0,238	-0,329	0,438	0,298	2	Secreted
381	P01717	Immunoglobulin lambda variable 3-25	0,616	0,124	0,545	0,941	2	Secreted
382	P0DOX8	Immunoglobulin lambda-1 light chain	-0,142	-0,493	-0,152	-0,545	1	Secreted
383	P15814	Immunoglobulin lambda-like polypeptide 1	-2,681	-2,945	-3,334	-2,337	1	Secreted
384	P0DOX6	Immunoglobulin mu heavy chain	-1,756	-3,282	-1,930	-1,188	1	Secreted
385	Q15181	Inorganic pyrophosphatase	-3,920	-4,049	-4,430	-3,916	1	Cytoplasm
386	P11215	Integrin alpha-M	3,103	2,818	2,521	2,590	15	Cell membrane
387	P05107	Integrin beta-2	2,213	2,042	1,719	1,969	7	Cell membrane
388	P19827	Inter-alpha-trypsin inhibitor heavy chain H1	-2,308	-2,208	-1,430	-0,232	2	Secreted
389	P19823	Inter-alpha-trypsin inhibitor heavy chain H2	-1,167	-1,314	-0,872	-0,431	4	Secreted
390	Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	0,233	0,452	1,001	1,955	7	Secreted
391	P18510	Interleukin-1 receptor antagonist protein	2,806	2,796	2,602	2,155	6	
392	Q14116	Interleukin-18	-0,397	-0,630	-0,642	-0,640	2	Secreted
393	Q9UHA7	Interleukin-36 alpha	0,585	0,922	0,441	0,508	3	Secreted
394	Q9UBH0	Interleukin-36 receptor antagonist protein	-2,914	-2,641	-2,561	-2,624	1	Secreted
395	P07476	Involucrin	-3,912	-3,584	-3,958	-3,797	2	Cytoplasm
396	P48735	Isocitrate dehydrogenase	-2,672	-1,717	-2,238	-1,833	1	Mitochondria
397	O75874	Isocitrate dehydrogenase [NADP] cytoplasmic	-2,022	-1,952	-2,053	-2,734	1	Cytoplasm
398	P14923	Junction plakoglobin	3,959	4,858	4,098	4,235	25	Cell junction
399	P06870	Kallikrein-1	3,637	2,406	3,026	3,520	9	
400	O43240	Kallikrein-10	-0,290	0,488	0,593	-0,007	2	Secreted
401	Q9UBX7	Kallikrein-11	0,259	-0,330	-0,116	0,261	1	
402	Q9UKR0	Kallikrein-12	-1,955	-0,428	-1,523	-1,030	1	Secreted
403	Q9UKR3	Kallikrein-13	1,443	1,855	1,742	1,339	5	Secreted
404	Q9P0G3	Kallikrein-14	-3,336	-2,792	-3,420	-3,157	1	Secreted
405	P49862	Kallikrein-7	-1,820	-1,590	-0,833	-1,630	1	Secreted
406	O76013	Keratin, type I cuticular Ha6	-0,700	1,049	-0,240	-0,330	5	
407	P13645	Keratin, type I cytoskeletal 10	5,032	5,862	5,539	5,150	24	
408	P13646	Keratin, type I cytoskeletal 13	7,772	7,529	7,236	7,075	36	Secreted
409	P02533	Keratin, type I cytoskeletal 14	4,381	5,327	5,124	5,358	16	Secreted
410	P19012	Keratin, type I cytoskeletal 15	-0,791	-0,819	-0,555	-0,687	4	Cytoplasm
411	P08779	Keratin, type I cytoskeletal 16	5,425	6,705	6,466	6,485	19	
412	Q04695	Keratin, type I cytoskeletal 17	2,610	4,371	3,803	4,159	16	
413	P08727	Keratin, type I cytoskeletal 19	1,824	1,784	1,547	1,615	10	Cytoplasm

414	P35527	Keratin, type I cytoskeletal 9	2,761	2,380	2,217	2,901	11	
415	Q9NSB2	Keratin, type II cuticular Hb4	0,616	2,476	0,776	0,854	12	
416	P04264	Keratin, type II cytoskeletal 1	5,315	5,915	5,584	5,389	16	
417	P35908	Keratin, type II cytoskeletal 2 epidermal	3,287	3,355	3,214	3,012	12	
418	Q01546	Keratin, type II cytoskeletal 2 oral	0,983	2,925	1,817	2,291	9	
419	P12035	Keratin, type II cytoskeletal 3	7,952	7,421	6,937	7,014	12	
420	P19013	Keratin, type II cytoskeletal 4	9,944	9,415	9,050	9,164	67	
421	P13647	Keratin, type II cytoskeletal 5	5,945	6,319	6,119	6,092	23	Cytoplasm
422	P02538	Keratin, type II cytoskeletal 6A	6,016	6,911	6,346	6,299	7	Cell membrane
423	P04259	Keratin, type II cytoskeletal 6B	1,564	2,113	2,023	1,947	2	
424	P48668	Keratin, type II cytoskeletal 6C	-2,943	-2,332	-2,231	-2,224	1	Cytoplasm
425	Q8N1N4	Keratin, type II cytoskeletal 78	5,633	5,669	5,229	5,336	31	
426	P05787	Keratin, type II cytoskeletal 8	-3,170	-3,180	-3,177	-3,755	1	
427	Q6KB66	Keratin, type II cytoskeletal 80	-0,918	-0,645	-0,856	-1,363	4	
428	P01042	Kininogen-1	-2,381	-2,601	-1,712	-0,614	3	Secreted
429	P22079	Lactoperoxidase	6,211	5,543	5,602	5,863	26	Secreted
430	P02788	Lactotransferrin	8,588	8,850	8,468	8,862	69	Secreted
431	P02750	Leucine-rich alpha-2-glycoprotein	-0,231	-1,281	-0,250	1,138	2	Secreted
432	P30740	Leukocyte elastase inhibitor	5,157	5,192	5,083	5,074	18	Secreted
433	P09960	Leukotriene A-4 hydrolase	1,272	1,071	0,975	1,231	8	Cytoplasm
434	P31025	Lipocalin-1	4,772	4,780	4,331	3,686	12	
435	P23141	Liver carboxylesterase 1	-2,697	-2,158	-2,952	-2,756	2	EPR lumen
436	P00338	L-lactate dehydrogenase A chain	4,305	4,488	4,427	4,466	10	Cytoplasm
437	P07195	L-lactate dehydrogenase B chain	-0,245	-0,262	-0,303	-0,225	4	Cytoplasm
438	O95274	Ly6/PLAUR domain-containing protein 3	0,750	0,581	0,422	0,739	5	Cell membrane
439	O95867	Lymphocyte antigen 6 complex locus protein G6c	0,129	1,020	0,159	-0,931	2	Cell membrane
440	Q14210	Lymphocyte antigen 6D	-1,480	-1,660	-1,453	-2,236	1	Cell membrane
441	Q7L5L3	Lysophospholipase D GDPD3	-1,808	-2,032	-1,652	-2,189	2	Membrane
442	P11279	Lysosome-associated membrane glycoprotein 1	-0,800	-0,473	-0,497	-0,362	2	Cell membrane
443	P61626	Lysozyme C	6,639	6,355	6,762	5,464	8	Secreted
444	P40121	Macrophage-capping protein	0,476	0,397	0,348	0,089	4	Nucleus
445	Q14764	Major vault protein	-2,060	-2,085	-1,965	-1,910	1	Cytoplasm
446	P40925	Malate dehydrogenase, cytoplasmic	0,662	0,929	0,718	0,891	4	Mitochondria
447	P40926	Malate dehydrogenase, mitochondrial	2,446	2,425	2,515	2,704	8	Cytoplasm
448	P08493	Matrix Gla protein	-5,935	-6,910	-6,806	-7,859	1	Secreted
449	P14780	Matrix metalloproteinase-9	3,384	3,068	2,704	3,241	10	Secreted
450	P01033	Metalloproteinase inhibitor 1	2,964	3,005	3,154	3,265	5	Secreted
451	O14880	Microsomal glutathione S-transferase 3	-1,730	-1,524	-2,143	-2,585	1	EPR membrane
452	P26038	Moesin	1,169	0,742	0,851	0,888	9	Cell membrane
453	P08571	Monocyte differentiation antigen CD14	-0,910	-1,168	-0,885	-0,841	3	Cell membrane

454	Q99685	Monoglyceride lipase	-1,395	-1,421	-1,647	-2,045	2	Cytoplasm
455	Q5SSG8	Mucin-21	-2,587	-2,726	-2,697	-2,128	1	Cell membrane
456	P98088	Mucin-5AC	-0,413	0,103	0,467	0,931	12	Secreted
457	Q9HC84	Mucin-5B	5,354	6,022	5,221	5,510	38	Secreted
458	Q8TAX7	Mucin-7	0,984	0,261	0,644	-0,335	2	Secreted
459	P24158	Myeloblastin	6,296	6,378	5,575	5,624	9	Cytoplasm
460	P41218	Myeloid cell nuclear differentiation antigen	2,092	2,148	1,406	1,387	5	Cytoplasm
461	P05164	Myeloperoxidase	6,465	6,257	5,918	6,012	22	Lysosome
462	P60660	Myosin light polypeptide 6	1,657	2,011	1,847	1,712	5	
463	P19105	Myosin regulatory light chain 12A	0,141	0,321	0,103	0,054	2	
464	P12882	Myosin-1	-1,275	-1,584	-1,070	-1,281	5	Cytoplasm
465	P13533	Myosin-6	-2,405	-2,814	-3,255	-3,310	1	Cytoplasm
466	P35579	Myosin-9	3,426	3,674	3,149	3,037	19	Cytoplasm
467	Q9UJ70	N-acetyl-D-glucosamine kinase	1,322	1,390	1,337	1,079	8	
468	P00387	NADH-cytochrome b5 reductase 3	-1,557	-1,319	-1,686	-1,681	2	EPR
469	E9PAV3	Nascent polypeptide-associated complex subunit alpha	-2,052	-1,280	-1,248	-1,855	2	Cytoplasm
470	Q14697	Neutral alpha-glucosidase AB	-1,418	-1,608	-1,965	-1,120	2	EPR
471	P22894	Neutrophil collagenase	0,826	0,572	0,682	1,088	5	Secreted
472	P59665	Neutrophil defensin 1	4,067	4,656	4,919	4,808	2	Secreted
473	P08246	Neutrophil elastase	6,585	6,731	5,987	6,197	13	Cytoplasmic vesicle
474	P80188	Neutrophil gelatinase-associated lipocalin	5,245	5,528	5,254	5,380	12	Secreted
475	P43490	Nicotinamide phosphoribosyltransferase	1,346	1,500	1,434	1,726	3	Nucleus
476	P29475	Nitric oxide synthase, brain	-3,608	-1,246	-2,341	-3,042	1	Cell membrane
477	P22307	Non-specific lipid-transfer protein	-2,714	-2,651	-2,589	-2,114	1	Cytoplasm
478	Q9C002	Normal mucosa of esophagus-specific gene 1 protein	-3,625	-2,729	-3,095	-3,109	1	Nucleus
479	P61916	NPC intracellular cholesterol transporter 2	-2,167	-2,247	-2,716	-2,261	1	Secreted
480	P80303	Nucleobindin-2	1,313	0,195	0,490	0,539	3	Golgi apparatus
481	P22392	Nucleoside diphosphate kinase B	2,194	2,373	2,200	2,078	5	Cytoplasm
482	Q6UX06	Olfactomedin-4	0,538	0,723	0,781	0,941	5	Secreted
483	Q92882	Osteoclast-stimulating factor 1	-3,245	-3,811	-3,443	-4,020	1	Cytoplasm
484	Q96FX8	p53 apoptosis effector related to PMP-22	-1,555	-1,111	-1,344	-1,717	1	Cell junction
485	O00151	PDZ and LIM domain protein 1	-2,877	-3,411	-3,669	-3,518	1	Cell membrane
486	O75594	Peptidoglycan recognition protein 1	1,106	1,100	0,686	0,955	3	Cytoplasm
487	P62937	Peptidyl-prolyl cis-trans isomerase A	2,912	3,105	3,055	2,861	8	Secreted
488	P23284	Peptidyl-prolyl cis-trans isomerase B	3,547	3,754	3,586	3,382	6	Cytoplasm
489	P62942	Peptidyl-prolyl cis-trans isomerase FKBP1A	-0,765	-2,430	-2,197	-1,457	1	EPR lumen
490	O60664	Perilipin-3	-1,689	-1,306	-1,114	-1,136	3	Cytoplasm

491	O60437	Periplakin	1,922	2,208	1,753	1,962	13	Cytoplasm
492	Q06830	Peroxiredoxin-1	1,975	2,330	2,061	2,040	5	Cell junction
493	P32119	Peroxiredoxin-2	2,294	2,654	2,375	2,991	7	Cytoplasm
494	Q13162	Peroxiredoxin-4	-3,802	-3,987	-3,407	-3,751	1	Cytoplasm
495	P30044	Peroxiredoxin-5	1,631	1,473	1,379	1,236	5	Mitochondria
496	P30041	Peroxiredoxin-6	3,137	3,417	3,016	2,985	11	Mitochondria
497	P51659	Peroxisomal multifunctional enzyme type 2	-1,621	-0,656	-0,988	-1,448	3	Cytoplasm
498	Q00325	Phosphate carrier protein, mitochondrial	0,308	0,696	0,282	0,419	2	Peroxisome
499	P30086	Phosphatidylethanolamine-binding protein 1	-0,232	-0,328	-0,077	-0,444	4	MIM
500	Q96G03	Phosphoglucomutase-2	-1,839	-1,601	-2,120	-2,209	2	Cytoplasm
501	P00558	Phosphoglycerate kinase 1	2,402	2,262	2,401	2,354	8	Cytoplasm
502	P18669	Phosphoglycerate mutase 1	1,396	1,319	1,284	1,458	5	Cytoplasm
503	P14555	Phospholipase A2	-1,133	-2,989	-2,551	-2,076	1	
504	Q6P4A8	Phospholipase B-like 1	-0,098	0,356	-0,157	0,053	2	Cell membrane
505	P55058	Phospholipid transfer protein	-0,063	-0,773	-0,620	-0,679	2	Lysosome
506	Q13835	Plakophilin-1	1,478	2,587	1,692	1,153	6	Secreted
507	Q9Y446	Plakophilin-3	-0,964	-0,941	-1,177	-0,859	3	
508	P05155	Plasma protease C1 inhibitor	0,748	1,274	1,920	2,824	5	Cell junction
509	P00747	Plasminogen	-2,197	-1,434	-0,653	-0,457	2	Secreted
510	P05120	Plasminogen activator inhibitor 2	-2,099	-2,043	-2,426	-1,411	1	Secreted
511	P13796	Plastin-2	3,872	3,492	3,233	3,515	16	Cytoplasm
512	P13797	Plastin-3	0,291	0,825	0,759	0,647	5	Cytoplasm
513	Q15365	Poly (rC)-binding protein 1	-2,588	-2,994	-2,020	-2,212	1	Cytoplasm
514	P21128	Poly (U)-specific endoribonuclease	-1,352	-1,480	-1,693	-1,446	1	Secreted
515	Q96GD3	Polycomb protein SCMH1	1,521	-0,271	1,090	-2,841	1	Cytoplasm
516	P01833	Polymeric immunoglobulin receptor	8,322	7,743	7,839	7,905	38	Nucleus
517	P26599	Polypyrimidine tract-binding protein 1	-2,269	-2,040	-2,476	-2,162	2	Cell membrane
518	P0CG39	POTE ankyrin domain family member J	-1,640	-2,809	-2,125	-2,048	1	Nucleus
519	P02545	Prelamin-A/C [Cleaved into: Lamin-A/C	2,617	2,727	2,642	2,567	17	Cytoplasm
520	P07737	Profilin-1	3,193	2,405	2,551	2,578	7	
521	Q8WUM4	Programmed cell death 6-interacting protein	0,545	0,799	0,444	0,527	5	Nucleus
522	O75340	Programmed cell death protein 6	-0,749	-0,531	-0,856	-0,930	2	Cytoplasm
523	P35232	Prohibitin	-2,521	-2,238	-1,854	-2,260	1	Cytoplasm
524	Q99623	Prohibitin-2	-0,072	0,161	0,070	0,074	4	EPR membrane
525	P12273	Prolactin-inducible protein	8,295	7,718	7,726	7,393	7	MIM
526	Q6MZM9	Proline-rich protein 27	-0,282	-2,646	-1,409	-1,327	1	MIM
527	Q14914	Prostaglandin reductase 1	-0,804	-1,381	-0,424	-0,612	2	Secreted
528	Q16651	Prostasin	0,532	-0,107	-0,011	-0,014	2	Secreted
529	Q06323	Proteasome activator complex subunit 1	-1,986	-1,541	-1,607	-1,609	1	Cytoplasm
530	Q9UL46	Proteasome activator complex subunit 2	-1,165	-0,772	-0,643	-0,567	2	

531	P25787	Proteasome subunit alpha type-2	-2,363	-2,445	-2,058	-3,795	1	
532	P25788	Proteasome subunit alpha type-3	-4,660	-4,221	-4,119	-4,533	1	
533	P25789	Proteasome subunit alpha type-4	-2,310	-1,327	-1,763	-1,523	1	Cytoplasm
534	P28066	Proteasome subunit alpha type-5	-1,571	-2,099	-1,596	-1,214	2	Cytoplasm
535	P60900	Proteasome subunit alpha type-6	-0,133	-0,300	-0,032	-0,230	3	Cytoplasm
536	O14818	Proteasome subunit alpha type-7	-3,214	-3,046	-3,414	-3,748	1	Cytoplasm
537	P49721	Proteasome subunit beta type-2	-0,659	-0,570	-0,447	-0,854	2	Cytoplasm
538	P49720	Proteasome subunit beta type-3	-3,622	-3,326	-3,889	-3,109	1	Cytoplasm
539	P28074	Proteasome subunit beta type-5	-3,360	-3,494	-3,512	-3,396	1	Cytoplasm
540	P28072	Proteasome subunit beta type-6	-1,239	-0,959	-0,780	-1,038	1	Cytoplasm
541	P28065	Proteasome subunit beta type-9	-3,408	-2,504	-2,913	-2,902	1	Cytoplasm
542	P02760	Protein AMBP	-0,941	-0,803	0,155	0,714	3	Cytoplasm
543	P07237	Protein disulfide-isomerase	4,197	4,144	4,236	4,295	25	Cytoplasm
544	P30101	Protein disulfide-isomerase A3	1,339	1,026	1,313	1,542	9	Secreted
545	Q15084	Protein disulfide-isomerase A6	-0,937	-1,095	-0,965	-0,974	2	EPR
546	Q96BQ1	Protein FAM3D	-0,456	-0,821	-0,982	-0,289	2	EPR
547	Q9NUQ9	Protein FAM49B	-2,541	-2,145	-2,260	-2,645	1	EPR lumen
548	Q6P5S2	Protein LEG1 homolog	4,753	3,409	3,892	3,532	10	Secreted
549	Q08AI8	Protein mab-21-like 4	-1,381	-0,656	-1,043	-1,408	2	Secreted
550	Q96TA1	Protein Niban 2	-2,753	-1,556	-2,047	-2,875	1	
551	Q8WVV4	Protein POF1B	-3,677	-1,480	-2,534	-2,667	1	Cytoplasm
552	P60903	Protein S100-A10	0,416	0,960	0,559	0,484	3	Cell junction
553	P31949	Protein S100-A11	2,833	3,305	2,884	2,907	5	
554	P80511	Protein S100-A12	2,402	2,583	2,011	2,183	6	Cytoplasm
555	Q9HCY8	Protein S100-A14	3,027	3,489	3,114	2,674	8	Secreted
556	Q96FQ6	Protein S100-A16	1,334	1,618	1,198	1,061	5	Cytoplasm
557	P29034	Protein S100-A2	1,631	2,555	2,170	2,280	4	Nucleus
558	P06703	Protein S100-A6	-4,800	-3,906	-4,970	-5,113	1	
559	P31151	Protein S100-A7	0,210	1,338	2,043	2,004	6	Cytoplasm; Cell membrane
560	P05109	Protein S100-A8	8,653	9,390	8,806	8,768	19	Cytoplasm
561	P06702	Protein S100-A9	3,631	4,127	3,834	3,646	7	Secreted
562	P25815	Protein S100-P	1,465	1,841	1,375	1,362	3	Secreted
563	Q99497	Protein/nucleic acid deglycase DJ-1	-1,996	-2,844	-1,860	-1,553	1	Nucleus
564	Q9UM07	Protein-arginine deiminase type-4	0,285	-0,626	-0,857	-0,335	3	Cytoplasm
565	Q08188	Protein-glutamine gamma-glutamyltransferase E	5,161	5,965	5,681	5,295	33	Cytoplasm
566	P22735	Protein-glutamine gamma-glutamyltransferase K	3,111	3,586	3,299	2,964	20	Membrane
567	Q04941	Proteolipid protein 2	-2,806	-1,657	-2,220	-2,061	1	Membrane
568	P00491	Purine nucleoside phosphorylase	-0,619	-0,280	-0,277	-0,342	3	Cytoplasm

569	P55786	Puromycin-sensitive aminopeptidase	1,040	1,012	0,684	0,970	5	Cytoplasm
570	A6NMY6	Putative annexin A2-like protein	-3,058	-3,480	-2,837	-3,169	1	Secreted
571	Q5VTE0	Putative elongation factor 1-alpha-like 3	4,278	4,715	4,461	4,509	14	Cytoplasm
572	O00764	Pyridoxal kinase	-3,867	-4,120	-4,050	-4,376	1	Cytoplasm
573	P14618	Pyruvate kinase PKM	2,982	2,925	3,095	2,882	16	Cytoplasm
574	P50395	Rab GDP dissociation inhibitor beta	1,148	1,247	1,460	1,240	6	Cytoplasm
575	P46940	Ras GTPase-activating-like protein IQGAP1	1,520	1,540	1,372	1,150	9	Cell membrane
576	P63000	Ras-related C3 botulinum toxin substrate 1	-1,410	-1,514	-1,673	-1,418	1	Cell membrane
577	P15153	Ras-related C3 botulinum toxin substrate 2	-0,935	-1,923	-1,863	-1,488	1	Cytoplasm
578	P61026	Ras-related protein Rab-10	1,825	2,054	1,873	1,897	5	Cytoplasmic vesicle
579	Q15907	Ras-related protein Rab-11B	2,077	2,375	2,126	2,035	6	Cytoplasmic vesicle
580	P61106	Ras-related protein Rab-14	-2,296	-2,215	-1,878	-2,169	2	Cytoplasmic vesicle
581	Q9NP72	Ras-related protein Rab-18	-2,874	-3,560	-2,830	-2,624	1	Cell membrane
582	Q9H0U4	Ras-related protein Rab-1B	-3,590	-3,519	-2,961	-3,840	2	Cytoplasm
583	P57735	Ras-related protein Rab-25	-2,339	-2,600	-1,793	-2,692	1	Cell membrane
584	P51159	Ras-related protein Rab-27A	1,042	0,983	0,479	0,780	2	Membrane
585	P61019	Ras-related protein Rab-2A	-0,592	-1,031	-0,573	-0,994	3	EPR-Golgi membrane
586	P51149	Ras-related protein Rab-7a	1,364	1,569	1,202	0,972	2	Cytoplasmic vesicle
587	P62834	Ras-related protein Rap-1A	1,306	1,319	1,118	1,053	3	Cell membrane
588	P63244	Receptor of activated protein C kinase 1	0,220	0,809	0,476	0,383	4	Cell membrane
589	Q9HD89	Resistin	0,038	0,412	-0,305	0,040	2	Secreted
590	O95197	Reticulon-3	-0,846	-0,714	-0,921	-1,191	1	EPR membrane
591	P00352	Retinal dehydrogenase 1	-3,903	-3,523	-3,135	-3,905	1	Cytoplasm
592	P02753	Retinol-binding protein 4	-2,438	-2,172	-1,660	-1,228	2	Secreted
593	Q53RT3	Retroviral-like aspartic protease 1	-2,234	-1,100	-1,404	-2,318	1	Membrane
594	P52566	Rho GDP-dissociation inhibitor 2	2,321	2,073	1,798	2,141	4	Cytoplasm
595	Q07960	Rho GTPase-activating protein 1	-1,183	-0,920	-1,158	-1,922	2	Cytoplasm
596	P34096	Ribonuclease 4	-0,668	-1,076	-1,157	-0,653	2	Secreted
597	P13489	Ribonuclease inhibitor	0,479	0,781	0,881	0,676	6	Cytoplasm
598	P02810	Salivary acidic proline-rich phosphoprotein 1/2	-0,520	-1,703	-1,317	-0,586	5	Secreted
599	P16615	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	-4,051	-3,544	-3,439	-3,660	1	EPR membrane
600	O95171	Sciellin	-1,947	-3,035	-3,166	-2,467	1	Cytoplasm
601	O95969	Secretoglobin family 1D member 2	-3,872	-4,316	-4,235	-3,699	1	Secreted
602	Q96QR1	Secretoglobin family 3A member 1	-3,878	-3,329	-3,849	-2,890	1	Secreted
603	P30153	Serine/threonine-protein phosphatase 2A	-3,633	-3,274	-3,293	-3,741	2	Cytoplasm

604	P67775	Serine/threonine-protein phosphatase 2A	-1,676	-1,894	-1,412	-1,649	2	Cytoplasm
605	P62140	Serine/threonine-protein phosphatase PP1-beta	-2,525	-2,378	-2,283	-2,363	2	Cytoplasm
606	P02787	Serotransferrin	5,547	5,748	6,400	7,045	62	Secreted
607	Q9UIV8	Serpin B13	3,178	3,742	3,427	3,355	11	Cytoplasm
608	P29508	Serpin B3	1,841	2,515	2,411	2,329	10	Cytoplasm
609	P48594	Serpin B4	-0,125	0,330	0,587	0,672	4	Cytoplasm
610	P36952	Serpin B5	2,547	3,120	2,743	2,727	6	Secreted
611	P35237	Serpin B6	0,190	0,592	0,141	0,079	2	Cytoplasm
612	P02768	Serum albumin	9,197	9,809	10,510	11,074	84	Secreted
613	P0DJI8	Serum amyloid A-1 protein	-1,878	-1,603	-1,060	-1,222	2	Secreted
614	P02743	Serum amyloid P-component	-1,180	-0,644	-0,509	0,384	4	Secreted
615	O75368	SH3 domain-binding glutamic acid-rich-like protein	-0,723	-1,283	-1,091	-1,521	2	
616	Q9UJC5	SH3 domain-binding glutamic acid-rich-like protein 2	-2,077	-2,187	-2,753	-2,831	1	Nucleus
617	Q9H299	SH3 domain-binding glutamic acid-rich-like protein 3	-0,887	-0,997	-0,924	-0,835	2	Cytoplasm
618	Q04837	Single-stranded DNA-binding protein	-4,795	-5,138	-4,954	-4,529	1	Mitochondria
619	P62314	Small nuclear ribonucleoprotein Sm D1	-1,148	-1,193	-1,567	-1,536	1	Cytoplasm
620	P22531	Small proline-rich protein 2E	-0,928	-0,217	-0,050	0,037	2	Cytoplasm
621	Q9UBC9	Small proline-rich protein 3	4,867	4,796	4,584	4,691	12	Cytoplasm
622	Q14515	SPARC-like protein 1	0,827	-0,838	-0,016	-0,336	3	Secreted
623	Q13838	Spliceosome RNA helicase DDX39B	-0,314	-0,435	-0,462	-0,431	2	Nucleus
624	Q7KZF4	Staphylococcal nuclease domain-containing protein 1	-2,515	-2,270	-1,847	-1,507	1	Cytoplasm
625	P38646	Stress-70 protein	0,346	-1,022	0,787	-2,222	2	Mitochondria
626	O00391	Sulfhydryl oxidase 1	-0,171	-0,218	-0,305	-0,271	4	
627	Q9Y6N5	Sulfide:quinoneoxidoreductase	-1,186	-0,290	-0,454	-0,081	3	Mitochondria
628	O00204	Sulfotransferase 2B1	0,189	0,851	0,137	0,321	4	Cytoplasm
629	P00441	Superoxide dismutase [Cu-Zn]	-3,986	-3,308	-4,062	-3,255	1	Cytoplasm
630	P04179	Superoxide dismutase [Mn]	-0,475	-0,476	-0,080	0,135	3	Mitochondria
631	Q6UWP8	Suprabasin	-3,647	-3,480	-3,870	-4,258	2	Secreted
632	Q99536	Synaptic vesicle membrane protein VAT-1 homolog	-4,085	-4,218	-3,638	-4,110	1	Cytoplasm
633	Q16563	Synaptophysin-like protein 1	0,177	0,514	-0,090	-0,089	1	Cytoplasmic vesicle
634	P78371	T-complex protein 1 subunit beta	-3,455	-3,982	-3,212	-3,350	2	Cytoplasm
635	P50991	T-complex protein 1 subunit delta	-0,623	-0,472	-0,231	-0,814	3	Cytoplasm
636	P48643	T-complex protein 1 subunit epsilon	-2,421	-2,601	-2,602	-3,070	2	Cytoplasm
637	Q99832	T-complex protein 1 subunit eta	-1,834	-1,067	-0,757	-1,259	1	Cytoplasm
638	P49368	T-complex protein 1 subunit gamma	-1,660	-1,716	-2,244	-1,666	1	Cytoplasm
639	P10599	Thioredoxin	2,674	2,800	2,498	2,228	7	Nucleus
640	Q9BRA2	Thioredoxin domain-containing protein 17	-1,307	-0,747	-1,240	-0,813	1	Cytoplasm
641	P26639	Threonine--tRNA ligase 1, cytoplasmic	-0,927	-1,408	-1,271	-1,120	2	Cytoplasm

642	P19971	Thymidine phosphorylase	-0,832	-0,332	0,050	0,464	4	
643	P37837	Transaldolase	3,638	3,659	3,542	3,553	9	Cytoplasm
644	P20061	Transcobalamin-1	3,862	3,190	3,454	3,444	11	Secreted
645	P61586	Transforming protein RhoA	0,708	0,772	0,719	0,479	3	Cell membrane
646	P37802	Transgelin-2	0,099	0,519	0,312	0,332	5	
647	P55072	Transitional endoplasmic reticulum ATPase	1,034	1,247	1,201	1,289	6	Cytoplasm
648	P29401	Transketolase	4,268	4,181	4,043	4,126	15	
649	P13693	Translationally-controlled tumour protein	0,357	0,687	0,559	0,255	3	Cytoplasm
650	P51571	Translocon-associated protein subunit delta	-3,886	-4,132	-3,327	-5,055	1	EPR membrane
651	P49755	Transmembrane emp24 domain-containing protein 10	-2,156	-1,602	-2,086	-1,987	1	Golgi apparatus
652	Q86T26	Transmembrane protease serine 11B	0,156	-0,204	-0,396	-0,310	3	Cell membrane
653	O60235	Transmembrane protease serine 11D	1,518	1,346	1,237	1,071	3	Cell membrane
654	Q9UL52	Transmembrane protease serine 11E	-0,044	-0,256	-0,497	-0,591	5	Cell membrane
655	P02766	Transthyretin	1,633	2,153	2,662	2,962	9	Secreted
656	P40939	Trifunctional enzyme subunit alpha	-2,021	-2,096	-2,078	-2,255	2	Mitochondria
657	P55084	Trifunctional enzyme subunit beta	-0,212	-1,280	-0,700	-1,546	1	Mitochondria
658	P60174	Triosephosphate isomerase	3,905	3,769	3,986	4,183	14	Cytoplasm
659	Q14134	Tripartite motif-containing protein 29	1,221	0,831	-0,009	0,836	4	Cytoplasm
660	O14773	Tripeptidyl-peptidase 1	-1,597	-2,274	-1,908	-1,395	1	Lysosome
661	P06753	Tropomyosin alpha-3 chain	0,081	-0,245	0,243	0,006	3	Cytoplasm
662	P07477	Trypsin-1	2,798	1,689	4,393	4,118	1	Secreted
663	P23381	Tryptophan--tRNA ligase, cytoplasmic	-3,641	-2,725	-2,472	-2,854	1	Cytoplasm
664	Q71U36	Tubulin alpha-1A chain	-1,229	-0,627	-1,062	-0,743	2	Cytoplasm
665	P68366	Tubulin alpha-4A chain	1,222	1,590	1,275	1,413	4	Cytoplasm
666	P07437	Tubulin beta chain	-1,641	-1,909	-1,669	-1,301	1	Cytoplasm
667	Q13885	Tubulin beta-2A chain	-1,600	-0,798	-1,699	-1,316	1	Cytoplasm
668	P68371	Tubulin beta-4B chain	-0,380	-0,333	-0,558	-0,339	2	Cytoplasm
669	Q9BW30	Tubulin polymerization-promoting protein family member 3	-3,411	-2,824	-3,087	-2,349	1	Cytoplasm
670	P09758	Tumour-associated calcium signal transducer 2	-1,482	-0,669	-1,091	-1,141	2	Membrane
671	P41240	Tyrosine-protein kinase CSK	-4,845	-3,993	-3,994	-4,088	1	Cytoplasm
672	P62979	Ubiquitin-40S ribosomal protein S27a	3,094	3,621	3,308	2,955	4	Ubiquitin
673	P61077	Ubiquitin-conjugating enzyme E2 D3	1,059	-0,405	-1,571	0,045	1	Cell membrane
674	P68036	Ubiquitin-conjugating enzyme E2 L3	-3,679	-4,410	-3,501	-3,662	1	Nucleus
675	Q13404	Ubiquitin-conjugating enzyme E2 variant 1	-0,797	-0,117	-0,279	-0,892	1	Nucleus
676	P30085	UMP-CMP kinase	-2,280	-2,340	-2,412	-2,663	1	Nucleus
677	Q16851	UTP--glucose-1-phosphate uridylyltransferase	-2,762	-2,717	-2,537	-3,082	2	Cytoplasm
678	Q96QK1	Vacuolar protein sorting-associated protein 35	-1,629	-1,249	-1,403	-0,695	2	Cytoplasm

679	P49748	Very long-chain specific acyl-CoA dehydrogenase	-1,062	-0,996	-1,178	-1,142	3	MIM
680	Q53GQ0	Very-long-chain 3-oxoacyl-CoA reductase	-3,482	-2,638	-3,482	-3,268	1	EPR membrane
681	P08670	Vimentin	1,822	0,790	0,710	1,092	7	Cytoplasm
682	P18206	Vinculin	-1,225	-2,506	-1,484	-2,269	2	Cell membrane
683	P02774	Vitamin D-binding protein	2,087	2,493	2,929	3,644	16	Secreted
684	P04004	Vitronectin	-0,548	0,032	0,764	1,093	3	Secreted
685	P21796	Voltage-dependent anion-selective channel protein 1	1,010	0,963	0,959	0,785	5	MOM
686	P45880	Voltage-dependent anion-selective channel protein 2	1,149	1,244	0,959	1,037	2	MOM
687	P38606	V-type proton ATPase catalytic subunit A	-2,686	-3,363	-2,518	-2,361	1	Cytoplasm
688	O75083	WD repeat-containing protein	1,558	1,762	1,604	1,703	5	Cytoplasm
689	O15231	Zinc finger protein 185	-3,519	-2,985	-3,008	-3,381	1	Cytoplasm
690	P25311	Zinc-alpha-2-glycoprotein	6,145	5,227	5,797	6,096	20	Secreted
691	Q96DA0	Zymogen granule protein 16 homolog B	9,866	9,504	9,270	8,519	10	Secreted

Table 17. List of differential proteins between control, PVL, early, and advanced OSCC groups. The UniProt protein ID and name are presented. The representative identification with the number (#) of peptides with which a protein was determined, the average log2 transformed quantification data (n=10) and Student's t-test *p*-value from the pair-wised analysis are given in each case.

			Log2Med				*p*-value			
Protein ID	Protein name	# peptides	Ctrl	PVL	eOSCC	advOSCC	Ctrl PVL	Ctrl eOSCC	Ctrl advOSCC	PVL OSCC
P01024	Complement C3	67	4,757	5,217	5,779	6,463	0,233	0,018	0,005	0,048
P02647	Apolipoprotein A-I	23	3,208	3,908	5,127	5,472	0,073	0,003	0,016	0,008
P00738	Haptoglobin	25	3,604	3,932	4,872	6,222	0,441	0,019	0,003	0,003
P02790	Hemopexin	14	2,913	3,464	4,081	5,105	0,174	0,013	0,003	0,027
P08670	Vimentin	7	1,822	0,790	0,710	1,092	0,009	0,029	0,012	0,779
P01859	Immunoglobulin heavy constant gamma 2	6	4,729	5,220	5,759	6,583	0,220	0,042	0,003	0,059
P31151	Protein S100-A7	6	0,210	1,338	2,043	2,004	0,055	0,001	0,001	0,184
P07737	Profilin 1	7	3,193	2,405	2,551	2,578	0,004	0,004	0,012	0,519
P01034	Cystatin C	9	3,147	1,569	1,930	2,007	0,001	0,028	0,032	0,331
P04040	Catalase	10	2,138	1,487	1,407	1,798	0,085	0,023	0,301	0,751
P01036	Cystatin S	2	3,679	1,155	0,485	1,460	0,005	0,002	0,074	0,836
P98088	Mucin -5AC	12	-0,413	0,103	0,467	0,931	0,172	0,023	0,049	0,183
P15104	Glutamine synthetase	6	1,099	2,033	2,056	1,765	0,000	0,000	0,015	0,578
Q8WVV4	Protein POF1B	1	-3,677	-1,480	-2,534	-2,667	0,001	0,077	0,157	0,078
P12532	Creatine kinase U-type	2	-0,349	-0,903	-0,969	-1,193	0,307	0,226	0,093	0,750
Q9Y6N5	Sulfide:quinone oxidoreductase	3	-1,186	-0,290	-0,454	-0,081	0,006	0,020	0,002	0,873
P02765	Alpha-2-HS-glycoprotein	3	0,474	0,993	1,639	2,101	0,006	0,020	0,002	0,873
P16615	Sarcoplasmic reticulum calcium ATPase 2	1	-4,051	-3,544	-3,439	-3,660	0,307	0,256	0,344	0,988
P18206	Vinculin	2	-1,225	-2,506	-1,484	-2,269	0,001	0,415	0,006	0,032
P48594	SERPIN B4	4	-0,125	0,330	0,587	0,672	0,051	0,003	0,016	0,189
P04004	Vitronectin	3	-0,548	0,032	0,764	1,093	0,250	0,020	0,035	0,127

Q04118	Parotid salivary glycoprotein G1	2	-1,161	-4,562	-3,513	-3,056	0,002	0,095	0,129	0,171
Q6MZM9	Proline-rich protein 27	1	-0,282	-2,646	-1,409	-1,327	0,018	0,203	0,276	0,077
P00367	Glutamate dehydrogenase 1	1	-2,634	-1,453	-1,781	-1,563	0,017	0,086	0,019	0,471
P05155	Plasma protease C1 inhibitor	5	0,748	1,274	1,920	2,824	0,204	0,005	0,003	0,025
P19971	Thymidine phosphorylase	4	-0,832	-0,332	0,050	0,464	0,033	0,010	0,000	0,023
Q9UM07	Protein-arginine deiminase type-4	3	0,285	-0,626	-0,857	-0,335	0,015	0,025	0,135	0,942
O95336	6-phosphogluconolactonase	3	-0,490	-1,498	-0,775	-0,923	0,002	0,149	0,044	0,028
P14555	Phospholipase A2	1	-1,133	-2,989	-2,551	-2,076	0,031	0,066	0,204	0,375
P31944	Caspase-14	2	-2,916	-1,439	-1,225	-2,766	0,020	0,009	0,775	0,205
P05090	Apolipoprotein D	3	-1,578	-0,754	-0,700	-0,466	0,109	0,057	0,040	0,704
P49411	Elongation factor Tu	1	-1,786	-3,586	-2,064	-3,465	0,001	0,590	0,001	0,103
P26583	High mobility group protein B2	2	0,088	-1,346	-1,497	-1,526	0,006	0,011	0,002	0,742
P19652	Alpha-1-acid glycoprotein 2	2	-0,258	0,121	0,947	1,289	0,409	0,003	0,005	0,038
P12830	Cadherin-1	2	-0,334	-0,644	-0,995	-0,807	0,088	0,026	0,021	0,173
P23381	Tryptophan--tRNA ligase	1	-3,641	-2,725	-2,472	-2,854	0,058	0,020	0,146	0,890
Q14515	SPARC-like protein 1	3	0,827	-0,838	-0,016	-0,336	0,000	0,065	0,008	0,046
A0A0C4DH31	Immunoglobulin heavy variable 1-18	1	2,150	2,107	2,370	2,393	0,908	0,531	0,584	0,322
P61254	60S ribosomal protein L26	2	-1,435	-0,662	-0,937	-1,270	0,090	0,274	0,760	0,092
O15511	Actin-related protein 2/3 complex subunit 5	1	0,857	0,623	0,356	0,175	0,202	0,004	0,000	0,048
P00747	Plasminogen	2	-2,197	-1,434	-0,653	-0,457	0,121	0,014	0,016	0,041
P36542	ATP synthase subunit gamma	2	-2,380	-1,852	-1,591	-1,694	0,375	0,189	0,162	0,645
P41240	Tyrosine-protein kinase	1	-4,845	-3,993	-3,994	-4,088	0,072	0,054	0,076	0,882
Q8IUE6	Histone H2A type 2-B	2	0,215	-0,958	-0,750	-0,272	0,064	0,178	0,268	0,437
O00303	Eukaryotic translation initiation factor 3 subunit	1	-4,550	-3,324	-4,027	-3,867	0,004	0,213	0,114	0,019
Q96GD3	Polycomb protein SCMH1	1	-0,765	-2,430	-2,197	-1,457	0,000	0,289	0,000	0,279
P62942	Peptidyl-prolyl cis-trans isomerase	1	-0,765	-2,430	-2,197	-1,457	0,000	0,008	0,073	0,166
P36957	Dihydrolipoyllysine-residue succinyltransferase	1	2,378	0,307	2,470	-0,944	0,000	0,814	0,000	0,370
P62857	40S ribosomal protein S28	1	3,068	0,291	-0,902	-1,023	0,065	0,037	0,005	0,426

XI. *ANNEX*

ETHICS COMMITTEE APPROVAL

AND

PARTICIPANTS' WRITTEN INFORMED CONSENT FORM

D. José María Montiel Company, Profesor Contratado Doctor Interino del departamento de Estomatología, y Secretario del Comité Ético de Investigación en Humanos de la Comisión de Ética en Investigación Experimental de la Universitat de València,

CERTIFICA:

Que el Comité Ético de Investigación en Humanos, en la reunión celebrada el día 12 de diciembre de 2016, una vez estudiado el proyecto de investigación titulado:
"Análisis de los elementos presentes en la respuesta inflamatoria, perfiles de glicanos y análisis proteómico como biomarcadores en el diagnóstico precoz del carcinoma oral de células escamosas", *número de procedimiento H1480794580696*,
cuyo responsable es D. José V. Bagán Sebastián,
ha acordado informar favorablemente el mismo dado que se respetan los principios fundamentales establecidos en la Declaración de Helsinki, en el Convenio del Consejo de Europa relativo a los derechos humanos y cumple los requisitos establecidos en la legislación española en el ámbito de la investigación biomédica, la protección de datos de carácter personal y la bioética.

Y para que conste, se firma el presente certificado en Valencia, a catorce de diciembre de dos mil dieciséis.

Av. Blasco Ibáñez, 13 tel. 9630 64109 vicerec.investigacio@uv.es
Valencia 46010 fax 9639 83221 www.uv.es/sennues

Título del proyecto: Análisis de los elementos presentes en la respuesta inflamatoria, perfiles de glicanos y análisis proteómico como biomarcadores en el diagnóstico precoz del carcinoma oral de células escamosas

INVESTIGADOR PRINCIPAL: Dr. José V. Bagán

OBJETIVOS:

Realizar un estudio de investigación para determinar si existen alteraciones en la saliva en los pacientes con precáncer y cáncer oral.

DESCRIPCIÓN DEL ESTUDIO

Se le va a realizar un estudio completo de la boca valorando las lesiones orales que tenga, mediante una inspección clínica.

Se tomará una muestra de la saliva en reposo del paciente sin ningún método de estimulación química o mecánica, antes de que le inicien cualquier tratamiento, para posterior análisis.

Si Ud. esta de acuerdo, libremente firme el consentimiento de participación en este estudio que para este fin se ha añadido al final de este impreso.

RIESGOS Y BENEFICIOS

No existen riesgos asociados a la toma de saliva.

Con su participación en este estudio, usted va a ayudar a conocer si existen algunas pruebas diagnósticas que ayuden al diagnóstico precoz del cáncer en la boca.

Esta información podrá ser aprovechada en su propia salud.

PARTICIPACIÓN EN EL ESTUDIO

Su participación en este estudio es totalmente voluntaria y no recibirá remuneración alguna.

Como paciente, el rechazo a participar no supondrá ninguna penalización o ni afectará en modo alguno a la calidad de la asistencia sanitaria que reciba.

CONFIDENCIALIDAD

Toda la información obtenida será confidencial, los datos recogidos se introducirán, por el Equipo investigador, en una base de datos para realizar el análisis estadístico pero su nombre no aparecerá en ningún documento del estudio, sólo se le asignará un número. En concreto, las muestras se identificarán con un número y se agruparan por patologías afines. En ningún caso se le identificará en las publicaciones que puedan realizarse con los resultados del estudio. Sin embargo, esta información podrá ser revisada por el Comité Ético de Investigación Clínica del Hospital General Universitario de Valencia así como por organismos gubernamentales competentes.

Puede ejercer su derecho de acceso y rectificación de sus datos. También, si así lo desea, puede ser informado de los resultados del estudio.

El estudio se realizará asegurando el cumplimiento de normas éticas y legales vigentes (Declaración de Helsinki).

Si tiene alguna duda o no entiende este texto consulte antes de firmar el documento con el Dr. José V. Bagán - teléfono 620224129 que es el médico responsable de esta investigación y le puede preguntar cualquier duda o problema que tenga relacionado con este estudio o consulte con sus familiares y, finalmente, si está de acuerdo firme este consentimiento. Se le entregará una copia.

Fdo.: José V. Bagán
Investigador Principal del Proyecto
Universidad de Valencia y Servicio de Estomatología y Cirugía Maxilofacial del Hospital General Universitario de Valencia (HGUV)
TEL: 620224129

Firma del participante:
Fecha: _____

CONSENTIMIENTO DEL PACIENTE SUJETO DE ESTUDIO

Título del estudio: Análisis de los elementos presentes en la respuesta inflamatoria, perfiles de glicanos y análisis proteómico como biomarcadores en el diagnóstico precoz del carcinoma oral de células escamosas

Yo,……..……

He leído la hoja de información anterior.

He podido hacer preguntas sobre el estudio.

He recibido suficiente información sobre el estudio.

He hablado con el Dr. José V. Bagán Sebastián y para la explicación del estudio.

Comprendo que mi participación es voluntaria.

Comprendo que puedo retirarme del estudio:

- Cuando quiera.
- Sin tener que dar explicaciones.
- Sin que esto repercuta en mis cuidados médicos.

Doy mi consentimiento para que este material aparezca en informes y artículos de revista de publicaciones médicas.

Entiendo que:

- Mi nombre no será publicado.
- El material no será utilizado para publicidad o embalaje.
- El material no será utilizado fuera de contexto.

Firmado…….. Fecha……..

XII. *RESUMEN*

RESUMEN

La incidencia de cáncer de cabeza y cuello (CCC) es superior a medio millón de casos al año en todo el mundo y representa aproximadamente el 3% de los tumoures malignos de adultos. El cáncer oral, sobre todo, el carcinoma oral de células escamosas (COCE), es la neoplasia maligna más común en la cavidad oral que cuando se diagnostica en estadios avanzados se caracteriza por un mal pronóstico y una baja tasa de supervivencia. A pesar de los avances terapéuticos y tecnológicos, el pronóstico para COCE no ha mejorado en las últimas décadas debido a que todavía muchos de ellos son diagnosticados en momentos tardíos, cuando los pacientes buscan atención al experimentar síntomas en su última etapa, etapa en la cual la enfermedad está avanzada y la tasa de supervivencia disminuye a niveles tan bajos como del 15-50%. Si por el contrario un carcinoma oral de células escamosas es diagnosticado de forma precoz o temprana, la supervivencia media a los 5 años llega a ser hasta del 85%. Tras la prevención primaria, el diagnóstico precoz se ha convertido en el principal objetivo en el manejo del cáncer oral. Por lo tanto, la detección temprana es esencial para mejorar las tasas de supervivencia y el pronóstico de estos pacientes. Sin embargo, las fases iniciales del cáncer oral son a menudo asintomáticas, siendo desde hace muchos años la biopsia la principal técnica para su diagnóstico. Actualmente, no existen biomarcadores para COCE con la sensibilidad y especificidad suficiente que nos permitan su uso clínico rutinario, lo que enfatiza la necesidad de alguna herramienta práctica y simple tanto para fines diagnósticos como de establecer programas de detección precoz o *screening*. Los tumoures malignos de la cavidad oral y faringe aparecen a través de una serie de mutaciones moleculares que resulten a un crecimiento celular descontrolado, pasando por distintas fases: desde momentos en los que se observan hiperplasias, lesiones displásicas, carcinoma *in situ* y, finalmente, pudiendo llegar a un auténtico carcinoma invasor. Existe un conjunto de alteraciones genéticas, epigenéticas y metabólicas que se han correlacionado con la transformación maligna en aquellos procesos potencialmente malignos en la cavidad oral. A las últimas pertenecen las lesiones y afecciones de la mucosa oral, con predisposición a la transformación maligna, según la Organización Mundial de la Salud (WHO). El desarrollo de muchos casos de COCE se ha correlacionado con la transformación maligna de lesiones orales potencialmente malignas, como la leucoplasia oral. Esta última presenta subtipos clínicos heterogéneos con diferentes potenciales de degeneración maligna. La leucoplasia verrugosa proliferativa (LVP) es con mucho la forma de leucoplasia con mayor riesgo de transformación

maligna. Los analitos salivales ya han demostrado su utilidad para la detección de trastornos locales, sistémicos e incluso en los procesos infecciosos. Los biomarcadores derivados de la saliva ofrecen un muestreo fácil con un mínimo riesgo para los pacientes, así como un reducido coste y tiempo de diagnóstico muy corto. Abarcan una variedad de parámetros detectables y medibles que permiten diferenciar la salud de la enfermedad. En términos de biomarcadores de respuesta inflamatoria, hay un aumento en la expresión de ciertas citoquinas. Se secretgan a partir de las células epiteliales e inmunes, lo que promueve la concentración de macrófagos y neutrófilos y produce inflamación. Como la inflamación se ha relacionado anteriormente con patologías orales, la tendencia actual en investigación indica la posibilidad de utilizar proteínas salivales pro y antiinflamatorias salivales para la detección de esas patologías, incluso en el caso del cáncer. Investigaciones recientes han demostrado el papel de ciertas citoquinas y su disregulación en el COCE, así como igualmente en lesiones orales potencialmente malignas (niveles salivales de factor de necrosis tumoural-α (TNF-α) y algunas interleucinas (IL), como la IL-1α, IL-6 e IL-8). Diversos estudios han presentado evidencias de que los niveles de citoquinas proinflamatorias salivales difieren significativamente entre individuos sanos, pacientes diagnosticados con COCE y aquellos con lesiones orales potencialmente malignas. La naturaleza local y sistémica de las respuestas inflamatorias alteradas sugiere que la expresión aberrante de citoquinas está relacionada con el precáncer y el cáncer y podría contribuir a la patogénesis de la neoplasia maligna en la cavidad oral. Los cambios en las modificaciones postraduccionales de proteínas (MPP) tienen un papel importante en el estudio de la etiología y la progresión de la enfermedad. La posibilidad de diferenciar los patrones de glicosilación de proteínas entre pacientes con cáncer y personas sanas demuestra como la glicobiología es un área potencial para el descubrimiento de biomarcadores. Se ha demostrado que la glicosilación aberrante acompaña a muchas afecciones crónicas e infecciosas y es una característica común de las células tumourales que puede afectar a los *N*- y *O*-glicanos en glicoproteínas, glicolípidos o glicosaminoglicanos. El perfil glicómico se ha utilizado recientemente para revelar cambios en los glicanos como resultado de la progresión de diferentes enfermedades malignas, incluyendo cáncer de mama, ovario, próstata y hepatocelular. Se ha descrito glicosilación anómala en el suero y tejidos malignos de pacientes con COCE, asociada con la invasión celular y la diseminación de la enfermedad. Sin embargo, faltan estudios sobre el potencial de los glicoconjugados salivales como biomarcadores del cáncer oral. Las alteraciones en la

glicosilación de proteínas pueden ser específicas de una patología, por lo que el estudio de los glicanos podría contribuir a una mejor comprensión de la carcinogénesis oral y proporcionar posibles herramientas para su diagnóstico. Se ha descrito y analizado, recientemente, el proteoma salival en los pacientes con cáncer de cabeza y cuello. Los estudios asociados a COCE han demostrado una expresión alterada de proteínas relacionadas con el metabolismo y la estructura celular, la adhesión y la motilidad, la transducción de señales, así como la inflamación, incluidas las oncoproteínas. La heterogeneidad celular y molecular de COCE y la gran cantidad de cambios moleculares involucrados en la carcinogénesis oral enfatizan la importancia de estudiar las proteínas responsables de ellas mediante proteómica de escala global.

Objetivos

El objetivo general de este estudio fue analizar un panel de citoquinas inflamatorias salivales, así como los perfiles proteicos y de glicanos salivales con el fin de identificar cuáles de ellas se detectarían de forma alteradamente significativos en las primeras etapas de cáncer oral y pudieran ser potenciales biomarcadores para el diagnóstico precoz de COCE.

Objetivos específicos:

1. Analizar si existen diferencias significativas en las citoquinas salivales (TNF-a, IL-1a, IL-6, IL-8, HCC-1, MCP-1 e PF-4), así como en los perfiles proteicos y glicanos salivales entre un grupo de pacientes con carcinoma oral de células escamosas (grupo 1), otro segundo con trastornos potencialmente malignos (grupo 2) y un grupo control (grupo 3).

2. Dentro del grupo con trastornos potencialmente malignos (grupo 2) se analizaron las diferencias en las citoquinas anteriores, así como en los perfiles proteicos y glicanos salivales entre los pacientes que presentaban dos formas clínicas de leucoplasia oral que eran la leucoplasia homógena (LH) y la leucoplasia verrugosa proliferativa (LVP).

3. Dentro del grupo con carcinoma oral de células escamosas se analizaron las diferencias en las citoquinas anteriores, así como en los perfiles proteicos y glicanos salivales entre los pacientes en etapas tempranas (estadios I y II) y aquellos con estadios avanzados (estadios III y IV) de COCE.

4. Finalmente, se estudiaron si habían diferencias significativas en los niveles de las citoquinas salivales, así como en los perfiles proteicos y glicanos salivales entre los diversos grupos de este proyecto de investigación.

Metodología

El estudio de las citoquinas inflamatorias consistió en un grupo 1 compuesto por 66 pacientes con carcinoma oral de células escamosa diagnosticados mediante biopsia, visitados en el Servicio de Estomatología de la Universidad de Valencia, en Medicina Bucal de la Clínica Odontológica de la Universidad de Valencia o en el Servicio de Otorrinolaringología del Hospital General Universitario de Valencia. Todos los pacientes fueron informados adecuadamente del contenido del estudio y firmaron el correspondiente consentimiento informado, según fue aprobado por el Comité de Ética del Hospital General Universitario de Valencia. Se determinó el estadio del COCE, considerando estadios iniciales el I y II. Los estadios tardíos fueron el III y IV. Todos los tumoures se clasificaron según el sistema de clasificación TNM (Tumour Node Metástasis). El grupo 2 estuvo constituido por 66 pacientes con leucoplasia oral, clínicamente diagnosticada (tanto las leucoplasias homogéneas como las leucoplasias verrugosas proliferativas) de acuerdo con los criterios de Van der Waal y basado en el informe histopatológico clínico de la biopsia de la lesión. A su vez tuvimos un grupo control (grupo 3) de 25 personas sanas, con edades y sexo similares a los de los grupos 1 y 2 y sin presentar trastornos patológicos en la mucosa oral ni en las glándulas salivales.

Las muestras de saliva se obtuvieron tras su diagnóstico y antes del tratamiento. Los especímenes salivales se recogieron mediante la técnica de expectoración siendo la saliva total sin estimulación, siguiendo el protocolo de Navazesh (1993). A los pacientes se les dieron las siguientes recomendaciones: los participantes no debían cepillarse los dientes dentro de los 45 minutos antes de la recogida de muestras, las muestras de saliva visiblemente contaminadas con sangre fueron descartados. Las muestras se centrifugaron a 3000 x g durante 15 minutos at 4°C para eliminar los gránulos. Luego se congelaron y almaceron a -80 ° C hasta su estudio molecular. En el caso de los grupos 2 y 3 las muestras salivales se obtuvieron con el mismo procedimiento descrito para el grupo 1. Se utilizaron estrategias ómicas combinadas para perfilar diferentes tipos de moléculas con potencial utilidad diagnóstica en la saliva de pacientes con LVP, estadios tempranos y avanzados de COCE así como en personas sanas. Para la detección y cuantificación de citoquinas salivales, se utilizó un inmunoensayo multiplexado para examinar los niveles de ocho proteínas inflamatorias diferentes. El formato más común y fue establecido para el ensayo multiplex basado en anticuerpos utilizando microesferas conjugadas con anticuerpos de corporación Luminex. La plataforma Luminex 200^{TM} para la detección de biomarcadores y análisis de proteínas consta

de inmunoensayos basados en microesferas magnéticas, instrumentos de detección y software. Los datos se analizaron mediante el software gestor de Bio-Plex. Las concentraciones de las ocho citoquinas se evaluaron por interpolación de la curva estándar usando una dilución de cinco veces por etapas de estándares de proteína. Las curvas de calibración se generaron para cada una con el software de gestión de Bio-Plex y las concentraciones se calcularon a partir de la curva estándar. Para evaluar las diferencias en los niveles de proteínas inflamatorias entre los individuos de los grupos de control, LH, LVP, pacientes con COCE inicial y avanzado, se realizó un análisis comparativo incluyendo la determinación del nivel de cuantificación, análisis de componentes principales para analizar la variabilidad y la prueba de chi-cuadrado intra e intergrupales para cada analito. Para valorar las diferencias en las medias de los niveles salivales de las citoquinas entre los grupos se aplicaron las pruebas no paramétricas de Mann-Whitney and Kruskal-Wallis. Se realizó un análisis de regresión y correlación entre los grupos y las citoquinas significativamente alterados para determinar si la expresión de las citoquinas mostraba una mayor relación con alguno de los grupos patológicos. Para evaluar el potencial diagnóstico de los biomarcadores candidatos descubiertos se construyeron curvas de característica operativa del receptor (ROC).

Para estudiar los perfiles de glicanos, se recogieron muestras de saliva de 20 individuos de la misma edad, igualmente distribuidos en cuatro grupos como se detalla a continuación: el grupo 1 consistió en 5 pacientes con LVP. El grupo 2 y 3 consistió de 5 casos cada uno, incluyendo pacientes en estadios iniciales y avanzados de COCE. El grupo 4- que fueron los controles, incluyo 5 personas sanas, sin visibles lesiones orales. Los *N*-glicanos liberados de las proteínas salivales se sometieron a una estimación analítica con cromatografía líquida de ultra alta resolución (UHPLC) acoplada a espectrometría de masas (MS). Todos los datos se adquirieron con el software Thermo Scientific Xcalibur 4.0. UHPLC se equipó con una columna de amida de glicano Waters BEH. Las adquisiciones de datos se proporcionaron utilizando Xcalibur (Thermo Fisher Scientific, MA, USA) software. El software de análisis de datos multivariados Progenesis QT (Waters TM) permitió la normalización de la abundancia de los glicanos y la cuantificación relativa entre los grupos estudiados. Las composiciones putativas de monosacáridos se determinaron manualmente utilizando la herramienta GlycoWorkbench 3.0. La interpretación de la relación entre los grupos experimentales se estimó mediante análisis de componentes principales (PCA), permitiendo visualizar datos multivariados. El

análisis estadístico y el diseño de gráficos se realizó con GraphPad Prism (GraphPad Software, Inc., San Diego, CA). Las diferencias entre grupos se calcularon mediante la prueba de Kruskal-Wallis.

El objetivo del estudio proteómico era identificar y cuantificar todas las proteínas que se pudieran detectar en las muestras analizadas en los diferentes grupos. La biblioteca espectral de proteínas salivales se creó mediante la identificación de proteínas presentes en una mezcla de muestras salivales de pacientes con COCE. Se seleccionaron 10 personas, incluyendo 3 casos de COCE inicial y 7 casos de COCE avanzados. La expresión diferencial de proteínas se estimó en las muestras individuales de la saliva obtenida en los 40 casos, distribuidas en cuatro grupos, de la siguiente manera: 10 controles, 10 LVP, 10 casos de COCE en estadios iniciales y 10 casos de COCE en estadios avanzados. LC-MS es una técnica de química analítica que combina las capacidades de separación física de la cromatografía líquida con las capacidades de análisis de masas mediante la espectrometría de masas. Mientras que LC separa muestras con múltiples componentes, MS proporciona identidad estructural y datos cuantitativos de los componentes individuales con alta especificidad molecular y sensibilidad de detección. Sequential Windowed Acquisition of All Theoretical Fragment Ions (SWATH)-MS es el método que se utilizó para la cuantificación de las proteínas presentes en las muestras salivales. Este enfoque analítico de alto rendimiento ha demostrado ser eficiente para los análisis de muestras complejas de una manera robusta y reproducible. Se utilizó el software ProteinPilot para generar una lista de picos basada en la medición de SWATH-MS. Los archivos obtenidos de la cuantificación SWATH fueron analizados por Peak View 2.1 con una biblioteca espectral de proteínas, construida previamente. La búsqueda de proteínas se realizó en la base de datos Swissprot. El análisis bioestadístico se llevó a cabo por métodos de regresión lineal y reducción de dimensiones (clasificación) a través del programa R 4.0.2, usando el paquete *glmnet* para LASSO (Least Absolute Shrinkage and Selection Operator) y Elastic Net Regularized Generalized Linear Models. Se utilizó la prueba t- test de student para determinar la significación estadística de las proteínas expresadas diferencialmente en base a la abundancia relativa dentro de las comparaciones de los grupos pareados estudiados. Las curvas ROC evaluaron el valor diagnóstico de los biomarcadores candidatos descritos. La versión 11 del online software STRING permitió buscar relaciones de interacción de las proteínas expresadas diferencialmente en el COCE comparados con los grupos de LVP y

controles. Las clasificaciones de los procesos biológicos hallados se realizaron con las herramientas de la base de datos STRING y, también se examinaron los términos de ontología genética enriquecida, proporcionados por el mismo software.

Resultados

Las medias de las concentraciones de cinco citoquinas, concretamente la IL-6, IL-8, MCP-1, TNF-α y HCC-1, fueron significativamente más altas en las leucoplasias (LH y LVP) que en el grupo de control. Además, los niveles de IL-6, IL-8, TNF-α, HCC-1 y PF-4 fueron considerablemente diferentes entre los pacientes con COCE en los dos estadios clínicos analizados comparados con los hallazgos en las personas sanas. Se observó una elevación notable de MCP-1 solo en las etapas iniciales de COCE, mientras que se hallaron niveles significativamente más altos de IL-6 y TNF-α en las etapas avanzadas en comparación con las iniciales de COCE. Se encontró que las medias de las concentraciones salivales de IL-6, IL-8, TNFα, HCC-1 y PF-4 seguían patrones de expresión similares con un aumento gradual desde los controles hasta los casos con leucoplasia (HL y/o PVL), y por supuesto mayor en los cánceres, desde los estadios iniciales a las etapas avanzadas de COCE. Las múltiples comparaciones mostraron que los niveles de IL-6, IL-8, TNF-α, HCC-1 y PF-4 estaban notablemente alterados en los grupos de pacientes con cáncer oral, mientras no se halló una alteración importante en las concentraciones de IL-1α, IP-10 y MCP-1 entre las diferentes patologías. Tampoco se observó una alteración significativa en los niveles estimados de citoquinas entre los grupos con lesiones de LH y PVL. Dado que los niveles de las seis citoquinas diferían notablemente entre los pacientes con COCE, tanto en los estadios iniciales y los controles, se analizaron las curvas ROC de estos biomarcadores. Los valores de AUC para IL-6, IL-8, TNF-α y HCC-1 salivales fueron superiores a 0.8 y para MCP-1 y PF-4 fueron superiores a 0.7. En una comparación de los niveles salivales de las citoquinas halladas entre los pacientes en etapas iniciales y avanzadas de COCE, la IL-6 y TNF-α mostraron concentraciones más altas, con valores de AUC de 0.69 y 0.68, respectivamente. Además, el análisis ROC reveló valores de AUC superiores a 0.7 para IL-6, TNF-α y PF-4 en las etapas iniciales del cáncer oral en comparación con los pacientes diagnosticados con leucoplasias. Se encontró una relación directa entre el grupo de COCE avanzado, con la TNF-α e IL-6, con una probabilidad de diagnosticar la enfermedad avanzada de hasta 70%. La presencia de metástasis en los ganglios linfáticos del cuello se asoció directamente con IL-6, que también

se correlacionó con los niveles de IP-10 y TNF-α. Según los rangos de concentración de IL-6, su probabilidad diagnóstica alcanza cifras de hasta 0.8 de 1.0. No se estableció correlación entre el sexo de los pacientes, el hábito de fumar, la ubicación de la lesión / tumour, la forma clínica y la expresión alterada de citoquinas.

La identificación y la descripción inicial de los *N*-glicanos liberados de las glicoproteínas salivales resultaron en una lista compuesta de 90 composiciones de monosacáridos. El perfil salival abundante de *N*-glucómico reveló la prevalencia de oligosacáridos neutros compuestos por estructuras complejas, híbridas y con un alto contenido de manosa. Los perfiles glicanos de abundancia relativa y representados por el gráfico del componente principales demostraron la segregación de los casos control con las patologías (LVP, estadios iniciales y avanzados de COCE). El análisis estadístico de los compuestos anotados reveló ocho *N*-glicanos que exhibían una abundancia relativamente significativa entre los grupos de COCE, LVP y controles. Tres de ellos demostraron una relativa disminución en la saliva de los pacientes con COCE en comparación con LVP y controles. Cinco casos representativos exhibieron una mayor abundancia relativa tanto en la saliva de los pacientes en los estadios iniciales, como en los estadios avanzados de COCE comparados con LVP y/o individuos sin lesiones orales (controles). Entre las estructuras de monosacáridos expresadas diferencialmente se observaron bi- y tri-antenarios *N*-glicanos con diversos grados de fucosilación y sialilación.

Un total de 1053 proteínas identificaron en las muestras salivales de COCE habiéndose incluido en la biblioteca de proteínas espectrales. De estas 1053 proteínas, 691 se identificaron con dos o más péptidos y una valor de confianza del 95% según la secuencia de lo péptidos detectados. Las alteraciones de los perfiles proteicos se estimaron mediante un análisis SWATH utilizado para la cuantificación de la abundancia relativa de proteínas individuales en las muestras de saliva de pacientes con COCE, LVP y personas sanas. El análisis de componentes principales y discriminante evaluó la capacidad de los perfiles proteicos para agruparse según el grupo de pertenencia. Este método analítico multivariante mostró la aglomeración de especímenes relacionados, dentro de los grupos definidos. El agrupamiento observado de las muestras de controles, LVP, estadios iniciales y avanzados de COCE indica una diferenciación y variabilidad entre los perfiles proteicos de los grupos investigados. Los datos que

representan la abundancia relativa de las proteínas se han obtenido por una transformación logarítmica y la interpretación de las muestras se ha hecho mediante Heatmaps. Tanto las muestras como las proteínas se representaron ordenados según el resultado de la clasificación jerárquica. Se realizó un análisis de regresión logística binomial para reducir el número de variables predictivas (proteínas) para delinear aquellas con un mayor potencial para discriminar entre los grupos. Para reducir el número de proteínas (variables predictivas) eliminando del modelo aquellas que no eran importantes, la penalización Lasso actúo como un método de selección de variables a través de la regresión logística. Este modelo se aplicó para el examen de grupos de dos por dos. El análisis destacó 49 proteínas con expresiones alteradas, clasificando las muestras según su grupo de pertenencia. La mayoría de estas proteínas diferenciales demostraron la capacidad de ordenar claramente las muestras así como el hecho de que fueran capaces de tener un valor discriminativo entre los grupos investigados. La significación estadística de la abundancia relativa de proteínas alteradas se validó por la prueba t-test de student con los valores *p* ajustados por FDR a partir de controles frente a pacientes con COCE en etapas iniciales y avanzadas. Para obtener más información sobre las modulaciones relacionadas con cáncer oral, se llevó a cabo una búsqueda exhaustiva de la literatura para comprobar cuál de las proteínas identificadas con expresión alterada se habían asociado con la oncogénesis. El resumen indicó que la mayoría de los marcadores que exhibieron expresiones diferenciales en el estudio actual ya estaban relacionados con el cáncer, incluido el COCE. También, se identificaron 23 proteínas presentes en la saliva de estos cánceres. Unas proteínas, que exhibieron una abundancia relativamente y significativamente alterada en el COCE según nuestros resultados y coincidiendo con los hallazgos de otros autores. Incluyen complement C3, alpha-1-antitrypsin (SERPINA1), cystatin-S (CST4), hemopexin (HPX), apolipoprotein AI (APOA1), plasminogen (PLG), protein S100-A7 (S100A7), profilin-1 (PFN1), etc. Nuestro estudio ha sido uno de los pocos que ha descrito perfiles salivales proteicos de pacientes con lesiones potencialmente malignas, como la leucoplasia oral y los ha comparado con aquellos pacientes con cáncer. Nuestra investigación se ha centrado especialmente en la LVP debido a su alto riesgo de transformación maligna. Los resultados destacaron varias proteínas expresadas diferencialmente que podrían discriminar entre las patologías de LVP y COCE. Algunos de ellos han sido descritos previamente en estudios relacionados con cáncer oral y lesiones asociadas. Endoplasmic reticulum chaperone BiP (HSPA5) y histone H2A type 2-A (HIST2H2AA) forman

parte de un panel de proteínas que se señaló su capacidad de ser capaces de detectar COCE y monitorizar lesiones con predisposición a la transformación maligna en el programa taiwanés de detección de cáncer en la cavidad oral. Los modelos predictivos para la sensibilidad y la sensibilidad de los biomarcadores identificados como candidatos se estimaron mediante curvas ROC y demostraron valores de AUC mayores de 0.7. Para resaltar los mecanismos moleculares potencialmente implicados en la alteración del microambiente de COCE, las proteínas expresadas diferencialmente se sometieron a exploración con la herramienta de anotación funcional STRING. El análisis de enriquecimiento mostró que el rango más alto de los procesos biológicos que involucran proteínas con expresión alterada en los grupos de COCE estaban relacionados con mecanismos reguladores activos que implican las respuestas inmunes e inflamatorias (APOA1, haptoglobin (HP), transthyretin (TTR), alpha-acid glycoprotein 1 (ORM1), catalasa (CAT), etc.). Además, la asociación más alta que encontramos en relación con la función molecular fue la inhibición de la actividad enzimática y la actividad de la proteína asociada al receptor (señalización) en las que los inhibidores de proteasa moleculares de la familia de proteínas serpins, cathepsin D (CTSD), proteína AMBP y CST4 se hallaron altamente relacionados.

Conclusiónes

1) Los hallazgos de elementos inflamatorios que incluyen IL-6, IL-8, TNF-α, MCP-1, HCC-1 y PF4 tanto en individuos sanos como en los pacientes con COCE pueden desempeñar un papel importante en la detección precoz del carcinoma oral de células escamosas. Se encontraron concentraciones salivales más altas de IL-6, IL-8, TNF-α, HCC-1 y PF4 en COCE que lo observado en las lesiones de HL y PVL.

2) La correlación de HCC-1 e IL-6 con las características histológicas de las lesiones de leucoplasia oral indica la posibilidad de que sirvan como marcadores de displasia epitelial. El TNF-α y la IL-6 pueden indicar la progresión del COCE. En el caso de la IL-6 esta puede ser de ayuda en la detección de adenopatías cervicales.

3) La IL-6 se sugiere como un potente biomarcador diagnóstico del COCE.

4) Las estrategias analíticas combinadas revelaron la composición de *N*-glicanos en la saliva de individuos sanos. El perfil comparativo de *N*-glicanos reveló varias estructuras bi- y tri-antenarios fucosilados, expresados diferencialmente entre los diversos grupos estudiados,

proporcionando una plataforma razonable para investigar más a fondo la utilidad de la glicosilación salival para el diagnóstico de COCE.

5) Los perfiles proteómicos generados por LC-MS incluyeron más de 600 proteínas cuantificadas en la saliva de personas sanas y con lesiones pre/cancerosas. El análisis comparativo dio como resultado una lista de proteínas expresadas diferencialmente, caracterizadas en COCE, que indicaban mecanismos significativamente alterados como el sistema inmunológico, la inhibición de actividades enzimáticas y la adhesión celular. Entre los marcadores candidatos identificados, algunos se habían descrito previamente en la saliva. Además, se han descrito varias proteínas nuevas asociadas al COCE.