UNESP - Universidade Estadual Paulista
"Júlio de Mesquita Filho"
Faculdade de Odontologia de Araraquara

Esteban Alexis Arroyo Ormeño

Busca de potenciais biomarcadores para leucoplasia verrucosa proliferativa

Araraquara

2022

UNESP - Universidade Estadual Paulista
"Júlio de Mesquita Filho"
Faculdade de Odontologia de Araraquara

Esteban Alexis Arroyo Ormeño

Busca de potenciais biomarcadores para leucoplasia verrucosa proliferativa

Tese apresentada à Universidade Estadual Paulista (Unesp), Faculdade de Odontologia, Araraquara para obtenção do título de Doutor em Ciências Odontológicas, na Área de Diagnóstico e Cirurgia

Orientadora: Profa. Dra. Andreia Bufalino
Coorientadora: Luciana Yamamoto Almeida

Araraquara

2022

ArroyoOrmeño, Esteban Alexis
A779b Busca de potenciais biomarcadores para leucoplasia verrucosa proliferativa / Esteban Alexis ArroyoOrmeño. -- Araraquara, 2022
175 p. : il., tabs.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Odontologia, Araraquara
Orientadora: Andreia Bufalino
Coorientadora: Luciana Yamamoto Almeida

1. Leucoplasia oral. 2. Proteômica. 3. Biomarcadores. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Odontologia, Araraquara. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Esteban Alexis Arroyo Ormeño

Busca de potenciais biomarcadores para leucoplasia verrucosa proliferativa.

Comissão julgadora

Para obtenção do título de Doutor em Ciências Odontológicas, na Área de Diagnóstico e Cirurgia

Presidente e Orientadora: Prof. Dra. Andreia Bufalino
Faculdade de Odontologia, Universidade Estadual Paulista (Unesp), Araraquara.

2º Examidora: Prof. Dra. Marisa Aparecida Cabrini Gabrielli
Faculdade de Odontologia, Universidade Estadual Paulista (Unesp), Araraquara.

3º Examinador: Dr. Tulio Morandin Ferrisse
Faculdade de Odontologia, Universidade Estadual Paulista (Unesp), Araraquara.

4º Examinadora: Dra. Susana Belen Bravo Lopez
Instituto de Investigación Sanitaria de Santiago de Compostela

5º Examinador: Prof. Dr. Mario Pérez Sayáns
Facultad de Medicina y Odontología, Universidad de Santiago de Compostela.

Araraquara, 20 de janeiro 2022

DADOS CURRICULARES

Esteban Alexis Arroyo Ormeño

NASCIMENTO: 12 de fevereiro 1992-Santiago-Chile

FILIAÇÃO: Sergio Antonio Arroyo Salas e Mireya Nayade Ormeño Hernández

FORMAÇÃO:

2018–2022: Doutorando em Ciências Odontológicas. Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, Sao Paulo, Brasil.
Orientador: Andreia Bufalino
Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

2017 – 2018: Diplomado em Pesquisa e Publicações em Saúde Pública. (Carga horária: 210h). Pontificia Universidade Católica do Chile, PUC, Santiago, Chile.

2017 – 2018: Estágio na área de Estomatologia (Carga horária: 352h). Hospital Barros Luco Trudeau, HBLT, Santiago, Chile

2017 – 2018: Estágio na área de Cirurgia Buco-Maxilo-Facial (Carga horária: 528h). Hospital Barros Luco Trudeau, HBLT, Santiago, Chile

2012 – 2017: Graduação em Odontología. Universidade do Chile, Santiago, Chile.

2010 – 2012: Bacharel com menção em Ciências Naturais e Exatas. Universidade do Chile, Santiago, Chile

Dedico este trabalho aos pacientes que gentil e voluntariamente participaram da pesquisa, pois sem vocês não seria possível a realização desse trabalho. Vocês são o principio e o fim, para que um dia o progresso da ciência possa continuar a oferecer a dignidade frente às doenças bucais graves e que tanto desolam o espírito de quem a padece.

AGRADECIMENTOS

Aos meus queridos pais, Sergio Arroyo e Mireya Ormeño, agradeço por tudo que me ensinaram e pelo tempo que se dedicaram para que eu tivesse a oportunidade de estudar e de ter não só uma boa formação profissional, mas também pessoal. Ensinaram-me a ser perseverante, disciplinado, respeitoso e agradecido. Os senhores sempre tiveram ao meu lado, não importando a distância que nos separava, dando-me força nos momentos de angústia, insegurança e tristeza. Eu dedico muito deste título aos senhores.

Aos meus amigos por acreditarem em mim, mesmo quando eu me desacreditava. Um especial agradecimento por terem feito parte desta luta, por terem me apoiado, pois sem vocês eu acho que não teria conseguido chegar nesse momento. De coração desejo que a vida nunca nos afaste totalmente uns dos outros.

À Professora Doutora Andreia Bufalino, minha orientadora, não só agradeço, como também declaro a minha mais profunda admiração e gratidão por ter me acolhido sempre como muita gentileza e me tirado da escuridão cognitiva em que eu estava submerso. Sem dúvida alguma afirmo que a sua presença no Programa de Doutorado foi um divisor de águas na minha vida profissional e pessoal. Não me esquecerei da oportunidade que a UNESP – Faculdade de Odontologia de Araraquara me proporcionou durante meu curso de Doutorado.

À CAPES:
O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) por meio do programa Estudantes-Convênio de Pós-Graduação – PEC-PG – Código de financiamento 88881.154490/2017-01.

“Nada na vida deve ser temido, somente compreendido. Agora é hora de compreender mais para temer menos". Marie Curie[*]

[*] D'Itri FM. Wastewater renovation and reuse: proceedings of the International Conference on the Renovation and Reuse of Wastewater Through Aquatic and Terrestrial Systems. New York: M. Dekker; 1977. p.iii.

Arroyo Ormeño EA. Busca de potenciais biomarcadores para leucoplasia verrucosa proliferativa [tese de doutorado]. Araraquara: Faculdade de Odontologia da UNESP; 2022.

RESUMO

O carcinoma espinocelular (CEC) representa mais de 95% de todas as neoplasias malignas que acometem a cavidade oral e muitas vezes estes tumores são precedidos por alterações clínicas que apresentam um evidente potencial de transformação maligna, as quais são chamadas de desordens potencialmente malignas orais (DPMO). Dentre estas, a leucoplasia oral (LO) é a mais importante das DPMOs com uma de incidência de 3,4% e uma taxa de transformação maligna que varia de 0,2% até 17,5%. Uma forma menos reconhecida e ainda pouco compreendida de leucoplasia, denominada leucoplasia verrucosa proliferativa (LVP), representa uma variante de comportamento persistente e progressivo para malignidade, com uma taxa de transformação maligna maior que 70%. O diagnóstico da LVP atualmente só é possível por meio da observação temporal e individual de cada paciente, com a demonstração de progressão clínica e histológica das lesões para um CEC. No entanto, ainda não existem métodos moleculares ou biomarcadores que possam de forma confiável auxiliar no diagnóstico diferencial e precoce entre LO e LVP. Além disto, a LVP frequentemente apresenta resposta inadequada a todas as modalidades de tratamento e sofre altas taxas de recidiva. Diante disto, identificar potenciais biomarcadores para LVP poderá auxiliar no diagnóstico diferencial, prognóstico e tratamento desta DPMO. Assim, a principal hipótese deste projeto é: O perfil proteômico entre LO e LVP é distinto e a sua caracterização poderá auxiliar no entendimento do comportamento clínico distinto entre estas DPMOs. Para testar esta hipótese, os objetivos específicos deste estudo consistem em (1) identificar potenciais biomarcadores por meio da associação da microdissecção a laser (ML) e espectrometria de massas em tandem (MS/MS); (2) avaliar possíveis mecanismos biológicos associados aos principais biomarcadores identificados por meio de ferramentas bioinformáticas e; (3) realizar estudos de validação *in vitro*.

Palavras chave: Leucoplasia oral. Proteômica. Biomarcadores.

Arroyo Ormeño EA. Researching of potential biomarkers for proliferative verrucous leukoplakia [tese de doutorado]. Araraquara: Faculdade de Odontologia da UNESP; 2022.

ABSTRACT

Oral squamous cell carcinoma (OSCC) represents more than 95% of all malignant neoplasms in the oral cavity and often these tumours are preceded by clinical entities which have a clear potential for malignant transformation, these are so-called oral potentially malignant disorders (OPMD). Among these, oral leukoplakia (OL) is the most prevalent OPMD with an incidence of 3.4% and a malignant transformation rate ranging from 0.2% to 17.5%. A less recognised and still poorly understood form of OL, is the proliferative verrucous leukoplakia (PVL), which represents a variant of recalcitrant and progressive behaviour towards malignancy, with a malignant transformation rate higher than 70%. Currently, the diagnosis of PVL is only based on the temporal and individual observation, with demonstration of clinical and histological progression of these lesions to OSCC. Regardless, there are still no molecular methods or biomarkers that can reliably assist in the differential and early diagnosis between OL and PVL. Moreover, PVL often shows an inadequate response to all treatment modalities and suffers high rates of recurrence. Therefore, identifying potential biomarkers for PLV may help in the differential diagnosis, prognosis and treatment of this OPMD. Thus, the main hypothesis of this research is: The proteomic profile between LO and LVP is different and its characterization may help in the understanding of the distinct clinical behaviour between these OPMD. To prove this, the main goals of this study are (1) identify potential biomarkers through the association of laser microdissection (LM) and mass spectrometry in tandem (MS/MS); (2) assess possible biological mechanisms associated with the main biomarkers founded by the aid of bioinformatics tools and; (3) performing *in vitro* validation assays.

Keywords: Oral leukoplakia. Proteomics. Biomarkers.

SUMÁRIO

1 INTRODUÇÃO

O carcinoma espinocelular (CEC) representa mais de 95% de todas as neoplasias malignas da boca[1,2], possui uma alta incidência e letalidade, apresentando 354.900 novos casos e 177.400 óbitos por ano aproximadamente[3]. Mesmo que existam avanços na compreensão e protocolos de tratamento para diferentes tumores malignos, as taxas de sobrevida em cinco anos para pacientes com CEC não excede os 50%-60%[1,4]. Dessa forma, a prevenção e o diagnóstico precoce são extremamente necessários para melhorar o prognóstico desses pacientes.

As desordens potencialmente malignas orais (DPMO) são condições clínicas que podem preceder ao CEC, e tem uma capacidade de transformação maligna variável. As DPMOs estão compostas por leucoplasia, leucoplasia verrucosa proliferativa, eritroplasia, eritroleucoplasia, líquen plano, estomatite nicotínica, queilite actínica, entre outras[2,5,6]. A leucoplasia oral (LO) é clinicamente definida como uma "placa branca com um risco questionável de malignidade, uma vez que todas as outras lesões clínicas semelhantes que não apresentam risco foram excluídas"[2,7], e apresenta taxas de transformação maligna dentre 2% a 8%[2,5,8,9]. Esta lesão pode apresentar aspectos microscópicos, que podem variar desde hiperceratose, com ou sem grau de displasia (baixo ou alto), até um carcinoma in situ[2,4,7].

Por outra parte, a Leucoplasia verrucosa proliferativa (LVP) é considerada de maneira geral como uma variante agressiva da LO[6,10–12]. Clinicamente, a LVP aparece como uma leucoplasia não homogênea que afeta locais únicos ou múltiplos, com envolvimento de áreas contíguas ou não contíguas ao longo do tempo, de crescimento lento e resistente a todas as formas de tratamento, além disso afeta principalmente a mulheres por sobre os 60 anos de idade, e sem presença de fatores de riscos como o fumo e álcool[13–17]. Histopatologicamente, a LVP pode variar desde hiperceratose com ou sem displasia a hiperplasia verrucosa, carcinoma verrucoso ou CEC, seja in situ ou infiltrante[10,18].

A LVP é a DPMO com a maior taxa de transformação maligna, recentemente foi determinada uma taxa cumulativa de 49,5% (IC99%: 26,7% - 72,4%) entre estudos com 12 a 20 anos de acompanhamento e uma taxa de transformação maligna por ano de 9,3%[8]. Além disso, devido ao alto potencial de recorrência e o comprometimento

vários sítios concomitantemente na cavidade bucal, não há tratamento disponível, mesmo antes do processo de transformação maligna[19–22]. Assim, o diagnóstico precoce da LVP é extremamente necessário para melhorar o prognóstico dos pacientes.

Por fim, o líquen plano oral (LPO) é uma doença inflamatória crônica de etiologia desconhecida, comumente caracterizada pela presença de lesões reticulares brancas bilaterais, em sua maioria simétricas[7,23–25]. Histologicamente é caracterizada pela presença de uma zona em forma de faixa bem definida de infiltração celular confinada à parte superficial do tecido conjuntivo, consistindo principalmente em linfócitos[23]. Atualmente, o LPO é considerado a DPMO com a menor taxa de transformação maligna[8], entretanto esta ideai é ainda controversa, em grande parte atribuível aos diferentes critérios diagnósticos empregados[2,8,24,25]. Além disso, alguns estudos têm demonstrado que a LVP, principalmente em estágios iniciais, pode apresentar aspectos clínicos e microscópicos que mimetizam o LPO[26–31]. Adicionalmente, desde o ponto de vista histológico a LVP apresenta um infiltrado inflamatório com predomínio de linfócitos T na lâmina própria, semelhante às características histopatológicas encontradas no LPO[23].

Deve-se destacar ainda, que a classificação proposta pela OMS enfatiza os critérios microscópicos para definir o estágio clínico das DPMOs, mas esses métodos representam um problema, pois são pouco reprodutíveis, têm alta variabilidade inter-observador e são limitados quando se trata de lesões com características histológicas semelhantes, principalmente para as lesões de LVP ou aquelas que podem ser confundidos com ela[2,12,32–35]. Apesar dos diferentes critérios desenvolvidos ao longo do tempo o diagnóstico da LPV ainda é controverso devido à falta de características clínicas, microscópicas e moleculares patognomônicas[10,18,36–38]. Contudo, há evidências recentes que reforçam a falta de consenso para o diagnóstico de LVP com base aos critérios histopatológicos[35], dessa maneira, ainda não há um método objetivo padronizado disponível para o diagnóstico da LVP assim como a sua detecção precoce, prevenção e reconhecimento da progressão tumoral a partir dessas lesões.

Além disso, não há dados sobre a caracterização das células imunes nas DPMO, as quais têm se mostrado como um fator contribuinte na transformação maligna, bem como na patogênese de diferentes tipos de tumor[39–43] e também como um provável

fator etiológico para o desenvolvimento do LPO, por meio da desregulação do sistema imune, envolvendo tanto a resposta inata como a adaptativa[44–46].

O uso de técnicas moleculares pode melhorar significativamente a detecção de alterações que são imperceptíveis na análise histopatológica, melhorando a precisão na identificação de pacientes com maior risco de desenvolver um carcinoma[14,47–49]. Ainda, as inovações no campo da identificação de proteínas que envolvem vários procedimentos altamente sensíveis como as técnicas de espectrometria de massas têm possibilitado melhorar a resolução da detecção de proteínas nos tecidos e mesmo até quando são aplicadas a fluidos complexos como a saliva[50–58].

Contudo, sabe-se que a saliva é uma matriz complexa, cuja composição pode variar em função da estimulação do sistema nervoso autônomo, ritmo circadiano, hábitos, estado de saúde-doença, entre outros[59–61]. Isso tem despertado grande interesse no uso da saliva como método de diagnóstico e controle de diversas doenças, tanto sistêmicas quanto da cavidade oral[60,61], devido ao abundante conteúdo molecular, menor invasividade e facilidade de extração, bem como o baixo custo envolvido[60]. Os avanços no campo da saliva estabeleceram sua utilidade como fonte de biomarcadores comparáveis ao sangue, líquido cefalorraquidiano, líquido pleural e urina[50,55,60,61], que permitiria a triagem precoce de diferentes doenças por meio de "biópsias líquidas"[55,56,61].

Por fim, diferentes softwares possibilitam a análise qualitativa e quantitativa do proteoma, permitindo explorar clusters e nós de diferentes redes, bem como suas vias biológicas, funcionais e metabólicas[52,57,58,62–64]. Entretanto, atualmente não existe um consenso sobre quais biomarcadores podem ser úteis para diagnosticar e/ou tratar a LVP[65].

Consequentemente, a busca por marcadores no tecido e na saliva é de suma importância, pois abre novas possibilidades de diagnóstico e possíveis alvos terapêuticos na LVP, o que levaria a melhorar a prevenção da progressão do carcinoma oral.

2 PROPOSIÇÃO

2.1 Objetivo Geral

Buscar biomarcadores que permitam diagnosticar e diferenciar a LVP das outras DMPO, independentemente das características histomorfológicas.

2.2 Objetivos Específicos

- Analisar comparativamente e caracterizar do infiltrado inflamatório da LVP e do LPO nas camadas epitelial e subepitelial.
- Procurar biomarcadores proteicos expressos na saliva já descritos na literatura por meio de uma meta-análise, para diferenciar as principais DPMO do tecido saudável e do CEC.
- Analisar descritivamente o proteoma da LVP em comparação com a LO em amostras de tecido e de saliva, buscar biomarcadores diagnósticos com potencial terapêutico, e com possibilidade de aplicação nas biópsias líquidas.

3 PUBLICAÇÕES

Os trabalhos científicos apresentados nesta secção foram desenvolvidos no decorrer do programa de doutorado graças a participação ativa do nosso grupo de pesquisa oriundo do departamento de Diagnóstico e Cirurgia da FOAr em parceria com os grupos Medicina Oral, Cirurgia Oral e Implantodontia e o Laboratório de Proteômica do Instituto de Investigação Sanitária de Santiago provenientes da Espanha. Adicionalmente, destacamos as contribuições de cada equipe de pesquisa das diferentes instituições brasileiras e finlandesas sinaladas em cada artigo especificamente.

3.1 Publicação 1*

Usefulness of protein-based salivary markers in the diagnosis of oral potentially malignant disorders: a systematic review and meta-analysis.

Esteban Arroyo[1], Sergio Piñeiro Donís[2], Cintia M Chamorro Petronacci[2,3], Monica G Oliveira Alves[4,5], Xabier Marichalar Mendía[6], Darcy Fernandes[1], Alejandro I Lorenzo Pouso[2], Andreia Bufalino[1], Susana Bravo López[7], **Mario Pérez Sayáns[1,2]***

The 1st and 2nd author have participated equally

Affiliations:

[1] Department of Diagnosis and Surgery, Araraquara, School of Dentistry, São Paulo State University (Unesp), Araraquara, SP, Brazil.

[2] Oral Medicine, Oral Surgery and Implantology Unit (MedOralRes). Faculty of Medicine and Dentistry, Universidade de Santiago de Compostela.

[3] Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain.

* O artigo segue as normas do periódico Cancer Biomarkers (ISSNs:1574-0153) no qual foi publicado.

[4] Anhembi Morumbi University, School of Medicine, São José dos Campos, São Paulo, Brazil.

[5] Universidade Mogi das Cruzes, School of Dentistry, Mogi das Cruzes, São Paulo, Brazil.

[6] Department of Stomatology II. School of Medicine and Dentistry. Leioa, Bizkaia.

[7] Proteomic Unit. Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain.

*****Correspondence**:** Prof. Dr. Mario Pérez-Sayáns PhD, MDS, FDS, DDS. Address: Faculty of Dentistry. Entrerríos s/n, Santiago de Compostela C.P. 15782 Spain. TLF: 0034626233504. E-mail: mario.perez@usc.es. https://orcid.org/0000-0003-2196-9868

Abstract

Using a meta-analytical approach this study aimed to analyse the diagnostic capacity of protein-based biomarkers in saliva for the differential diagnosis of oral potentially malignant disorders (OPMDs) and oral squamous cell carcinoma (OSCC) and a control group of healthy individuals (HCG).

Articles on protein-based biomarkers in saliva providing quantitative expression in individuals with clinically and histopathological diagnosed of OPMD or oral leukoplakia (OL) were considered eligible. Searches were conducted in eight electronic databases. The methodological quality was assessed through the Quality Assessment of Diagnostic Studies tool (QUADAS-2). Functional analysis was also performed. Meta-analyses were performed using the OpenMeta tool (Analyst).

Meta-analysis was possible for 4 of the 11 biomarkers studied. Only carcinoembryonic antigen (CEA) and the soluble fragment of cytokeratin 19 (CYFRA21) were significant for the OSCC/OPMD subgroup, both with very low heterogeneity. CEA had an

OE=25.854 (CI95%: 13.215-38.492, p<0.001, I2=0) and CYFRA21 an OE=9.317 (CI95%: 9.014-9.619, p<0.001, I2=0). For the OPMD/HCG subgroup, only CYFRA21 was significant, with an OE=3.679 (CI95%: 0.663-6.696, p=0.017) although with high heterogeneity (I2=91.24).

The CEA and CYFRA21 markers had proved to be very useful when differentiating OSCC from OPMD. The CYFRA21 resulted to be the only protein capable of distinguishing between OPMD and healthy controls.

Keywords: Systematic review; meta-analysis; protein-based biomarkers; saliva; oral leukoplakia; oral potentially malignant lesions.

1. Introduction

The oral squamous cell carcinoma (OSCC) represents more than 95% of all malignant neoplasms of the mouth[1,2], and it has a high prevalence and morbidity, with 354,900 new cases and 177,400 deaths registered per year worldwide[3]. Despite advances in the understanding and treatment protocols for different malignant tumours, the five-year survival rate for patients with OSCC does not exceed 50-60% overall[1,4]. As a result, prevention and early diagnosis are considered to be highly necessary to improve the prognosis and survival rates.

The oral potentially malignant disorders (OPMD) are clinical conditions that have a variable malignant transformation capacity. This group of lesions is comprised of leukoplakia, erythroplakia, erythroleukoplakia, lichen planus, nicotinic stomatitis, actinic cheilitis, amongst others[2,5,6]. The oral leukoplakia (OL) is the main lesion within the OPMD group, and has an annual global average of malignant transformation rate between 1 and 2%[2,5,7], nonetheless, these estimations may be higher, depending

on the geographical region observed and the level of dysplasia, meaning, that it is possible to achieve rates that are equal to or greater than 39%[2]. The OL is clinically defined as a "white plaque with a questionable risk of malignancy once all other similar clinical lesions that carry no risk have been excluded"[2,8]. This lesion may present microscopic aspects, which can vary from hyperkeratosis, with or without various levels of dysplasia (low or high), to carcinoma *in situ*[2,4,8]. Therefore, a biopsy is strictly necessary to establish a definitive diagnosis.

Currently, the diagnoses that are based on these clinical and histopathological criteria represent a problem since they are poorly reproducible and have high inter-observer variability, therefore impeding a real standardisation in the early detection and prevention of the development of OSCC[2,9–11]. As a result, one of the most important challenges in the diagnosis of OL is to predict which of the lesions will progress to OSCC.

Saliva is a complex biological matrix, the composition of which can vary depending on multiple factors[12–14]. The idea of using saliva as a method for the diagnosis and control of different diseases (periodontitis, cancer, recurrent aphthous stomatitis, OPMD, etc.)[13,14] has proved to be of great interest, not only due to the varying molecular content, the low invasiveness of the procedure and the fact that it is easy to collect, but also given the low-cost involved[13]. The advances that have been made into the study of saliva have meant that it has been possible to establish its usefulness as a source of biomarkers that is comparable to blood, cerebrospinal fluid, pleural fluid and urine[13–16], and which allows for the early diagnosis and screening of different diseases through *"liquid biopsies"* [14,15,17].

Despite innovations in the field of protein identification that involve several highly sensitive procedures (such as mass spectrometry techniques)[15,16,18–20], the ELISA method is still one of the most widely used, given its wide dissemination, reproducibility and low cost in comparison with the new omics techniques[21–24]. As a consequence, the search for protein markers in saliva offers further possibilities for understanding the structural interrelationships of OPMD and how these might determine progression to oral carcinoma.

The main goal was to perform a systematic review and meta-analysis with the markers that met the established requirements. The secondary goals were 1) To obtain information on the protein-based salivary markers described in the literature for OPMD, 2) To study the expression differences in comparison with healthy controls and malignant OSCC lesions.

2. Methods

A systematic review of the literature was conducted by a multidisciplinary team between the 1st of December, 2019 and the 31st of January, 2020. The review was written following the Cochrane Handbook for Systematic Reviews of Interventions version 5.1.0[25] and the PRISMA guide checklist[26,27] for systematic reviews (Supplementary 1).

2.1. *PICO question*

The study was organised based on the following PICO question: Patient, index test, comparison, outcome. P: patients diagnosed with an OPMD or OL. I: protein biomarker in saliva. C: patients with OSCC and control group of healthy individuals (HCG).. O: Expression differences between OPMD and OSCC-HCG. With this we generated the scientific question: Are there any protein-based markers in saliva that

make it possible to quantitatively obtain diagnostic differences between healthy patients, OPMD and OSCC?

2.2. Criteria for considering studies for this review

2.2.1. Types of accuracy study

Studies of protein-based molecular markers *(index tests)* in saliva that revealed quantitative expression results related to clinically and histopathologically diagnosed diseases *(reference standard)*. Therefore, the following were included 1) Human studies and controls, in the English language, considering the OL or the OPMD that included cases of OL as the *reference standard*; 2) Studies that included protein and saliva-based markers exclusively; 3) Studies that used ELISA and/or equivalent methods of analysis with quantitative expression determination (Bromocresol green method, Immunoturbidometry and LUMINEX). The exclusion criteria included: 1) Patients with explicit diagnosis, systemic disease or a disorder that interfered with the conditions of interest; 2) Studies that only evaluated blood or tissue markers; 3) Studies in which the biomarker was not validated and was only discovered in label-free analysis; 4) Studies that only tried to validate a technique for diagnosis; 5) Betel quid chewers population as the only risk factor; 6) Studies that did not define the biomarker; 7) Studies that investigated changes in cytokines in saliva pre/post-treatment and that did not distinguish a potential biomarker; 8) Studies that did not define or include groups with the conditions of interest; 9) Studies that only included the OSCC or OSCC/HCG group. 10) Studies that quantified protein expression in units of measurement other than weight/capacity (i.e. mg/dl, ng/dl or pg./dl), such as Sequence coverage, mass/load ratio (m/z), Ion score, Fold-change o Mascot score. 11) Studies in which the marker was expressed under a minimum breakpoint.

2.2.2. Participants

The participants included in the study were those with clear and validated diagnoses both clinically and histopathologically. Studies on experimental animals or in vitro models were excluded.

2.2.3. Control and target conditions

The patients included had been diagnosed according to the WHO Classification for Head and Neck Tumours. OPMD and OL were considered as the *target condition* and OSCC and HCG as the *healthy control group*, both positive and negative according to the compared subgroups. The HCG group included patients without systemic diseases and who were not treated with immunosuppressive medication.

2.2.4. Reference standard

The reference standard for the diagnosis of OPMD, OL and OSCC was defined based on clinical and histopathological parameters (dysplasia and malignancy). Studies that did not detail any reference standard for the diagnosis of the described pathologies were not considered eligible for inclusion in this review. Given the great diversity of studies, clear inclusive diagnostic criteria were established. For the OSCC group: clinically diagnosed patients who had been verified through a histopathological study, regardless of the degree of differentiation, excluding cases of warty carcinomas. For the OPMD group: cases of oral lichen planus, erythroplakia, oral submucous fibrosis and OL were included. We considered as OPMD any studies that did not mention a clear clinical diagnosis, but which mentioned non-malignant lesion with some degree of dysplasia. [28,29]. When was found solely the OL as the OPMD group, this must have been: white, non-detachable, homogeneous and non-homogeneous (mottled or

warty) lesions with hyperkeratosis, acanthosis and/or histological squamous hyperplasia, with any degree of dysplasia (except carcinoma in situ which was excluded).

2.2.5. Index test(s)

Any single protein-based biomarker detected in saliva, which was analysed from a quantitative point of view to differentiate between controls and targets. Studies analysing multiple biomarker panels were excluded because of the impossibility of analysis.

2.3. Search methods for the identification and selection of studies

2.3.1. Information sources and search strategy

Electronic searches in MEDLINE were performed using PubMed, EMBASE, OVID, Web of Science, Scopus, Cochrane Library, ClinicalTrials and the WHO regional bibliographic databases (AIM, IMEMR, IMSEAR, WPRIM) and the Conference Citation Index databases. Following the recommendations established by the Cochrane Systematic Reviews of Diagnostic Test Accuracy Group [30], the search strategy used to identify precision studies included three sets of terms related to the target condition (OPMD or OL), the index tests (protein-based biomarkers) and the type of oral sample (saliva) (Supplementary 2).

2.4. Data collection and analysis

The search was conducted by 2 observers (MPS and EA). Any ongoing discrepancies were resolved by discussion with 2 different reviewers (XMM and CMCP.). The reasons for excluding studies were detailed. During the first phase, the title and abstract of the retrieved articles were read, and any studies that complied with the inclusion

criteria, as well as those that did not provide sufficient data for a clear decision to be made were judged in full text. Subsequently, the entire content of all of the studies that were considered eligible was reviewed in their entirety in a second-round and a final decision was made as to whether they were to be included or not.

2.5. Assessment of methodological quality

Two reviewers (MPS and EA) assessed the quality of the included studies separately using the revised Quality Assessment of Diagnostic Studies tool (QUADAS - 2) [31]. This tool consists of four domains: 1) patient selection, 2) index test, 3) reference test, 4) flow and time. Besides, each domain was evaluated in terms of its risk of bias and the first three domains were also evaluated in terms of their applicability.

2.6. Qualitative analysis

The unit of analysis was the expression of each protein biomarker in saliva, supplemented by the author and the year of publication. In most studies, the quantitative expression or level of a quantifiable protein was reported. Accordingly, these results were interpreted as the presence or absence of disease (OSCC, OPMD or OL) based on a numerical measurement that was classified according to a specific threshold (predetermined or not). The expression estimates were shown as integer values and their standard deviation for each biomarker and subgroup. The weighted mean difference (WM) between the case group and the control group was obtained for each model, giving the overall expression (OE) and the 95% confidence interval (CI95%). We distributed the meta-analysis in two subgroups, the first OSCC (standard reference)/ OPMD and the second OPMD (standard reference)/HCG, including the isolated OL reports within OPMD group.

2.7. Quantitative analysis

A meta-analysis was performed when the number of data for a biomarker was at least two papers and when the units of measurement of the marker were the same. The analysis strategy was to include all expression studies, regardless of their threshold value. The statistical analysis of the results was performed using the OpenMeta tool (Analyst).

2.8. Functional networks, and pathway mapping

Analyses of biological pathways within the differential protein markers were performed using the PANTHER enrichment analysis tool [32], only for interactions with high confidence (0.7). The Reactome Biological Pathway enrichment analysis of proteins of interest was performed using the STRING tool [33]. In both cases, we used proteins that presented statistically significant differences in our study, and in some cases, we performed node enrichment to obtain more protein interactions. Finally, we used a systematic bioinformatics analysis including Database for Annotation, Visualisation and Integrated Discovery (DAVID) and gene ontology enrichment [34].

2.9. Assessment of heterogeneity

Firstly, the threshold effect of expression was evaluated graphically by looking at the forest plot and secondly, the threshold effect was evaluated statistically using the I^2 test. The p values>0.1 and an index I2<50% indicated a low heterogeneity among the studies and as such, the fixed-effect model was performed. On the other hand, p values<0.1 and an index I2>50% indicated considerable heterogeneity. All of the bilateral differences with a p-value less than or equal to 0.05 were considered as significant [35].

3. Results

3.1. Study Selection

The overall combined result of the references obtained from the aforementioned databases was 279 citations. After removing the duplicates, we obtained 217 unique citations. The biomarkers described in the literature for OPMD included: carcinoembryonic antigen (CEA), C-reactive protein (CRP), soluble fragment of cytokeratin 19 (CYFRA21), human epidermal growth factor receptor 2 (Her-2/neu; erbB-2), interleukin-1α (IL-1α), interleukin-1B (IL-1B), interleukin-6 (IL-6), interleukin-8 (IL-8), Naa10p, Resistin (RETN), Tumour necrosis factor-alpha (TNF-α). Out of these papers, only eight[36–43] fulfilled all of the eligibility criteria. The flowchart of research and data processing is shown in Fig. 1.

3.2. Characteristics of protein-based studies in saliva

Table 1 summarises the characteristics of the eight selected studies, disaggregated by biomarkers. The 8 references were conducted in 5 different countries: China[36], Italy[43], India[37,39–41], USA[42] and Taiwan[38]. A total of 986 patients were included who were distributed in the following groups: HCG (N=335), OL (N=20), OPMD (N=267) and OSCC (N=364). The average age of the HCG was 49.9 years, and it was 50 years in the OL group, 49.6 years in the OPMD and 55.9 years in the OSCC group. The analysed markers were: CEA (GO:0006915, apoptotic process), CYFRA21 (GO:0006955, immune response) IL-6 (GO:0006955, immune response) and IL-8 (GO:0006955, immune response). With regards to the level of dysplasia, this aspect of OPMD was only described in the studies by Rhodus et al[42] (Moderate (38,46 %), Severe (61,53)); Punyani et al[40](Mild (20%)) and Sharma et al[41](Mild (40%), Moderate (35%), Severe (25%)). No information was provided with regards to the size or clinical appearance of the OL (homogeneous non-homogeneous), or the treatment modality or follow-up (oncological events or recurrences).

3.3. Quality assessment of protein-based studies in saliva

Regarding the methodological quality of the studies (Supplementary 3), after applying the QUADAS-2 scale regarding the risk of bias, it was determined that this was very high for *patient selection* (100%) and very low (0%) for *Flow and Timing* and *Reference Standard*. Respect to *concerns regarding the applicability*, the methodological risk was generally very low, especially for the *patient selection* and index test fields, although the study by Zheng, J. et al (1 study of 8, 12.5 %) had an uncertain evaluation of the *reference standard*, given that it did not define the number of cases for each type of diagnosis included in the OPMD group.

3.4. Synthesis of the qualitative and quantitative analyses of four salivary biomarkers

In terms of the different markers and study subgroups, Table 2 shows all the values of OE and I^2. Only two markers were significant for the OSCC/OPMD subgroup, CEA and CYFRA21, both with very low heterogeneity. CEA had an OE=25.854 (CI95%: 13.215-38.492, $p<0.001$, $I^2=0$) and CYFRA21 an OE=9.317 (CI95%: 9.014-9.619, $p<0.001$, $I^2=0$). Both markers included two studies respectively (CEA: Zheng et al and Airoldi et al and CYFRA21: Awatashi et al and Rajkumar et al) (Fig. 2A and 2B). IL-8 was not a predictable marker for the OSCC/OPMD subgroup (Fig. 2C). For the OPMD/HCG subgroup, only CYFRA21, with an OE=3.679 (CI95%: 0.663-6.696, $p=0.017$) although with a high heterogeneity ($I^2=91.24$) (Table 2, Fig. 3A). CEA showed differences between OPMD and HCG approaching significance ($p=0.061$) but with high heterogeneity, additionally, IL-6 and IL-8 showed no differences and presented extreme ($I^2=99.89$) and low ($I^2=39.30$) heterogeneity, respectively (Fig. 3B, 3C, 3D).

Once these proteins had been obtained from the 8 selected references, we performed interaction studies to determine the possible relationship between them. To do this we used the STRING program, and as can be seen in Fig. 4, there is a direct relationship between IL-6 and IL-8 (CXCL-8), forming a small cluster of pro-inflammatory cytokines with two proteins, the CXC chemokine receptor type 2 (CXCR2) and the IL-6 receptor (IL6R), which are potent chemotherapeutics targets and neutrophil activators (Supplementary 4).

Given the relationship between the cytokines, which showed a well-defined cluster, several studies were performed to see the position of these proteins in the immune system's response to the inflammatory process. Using the DAVID program, we were able to observe how these proteins are found in the TNF signalling pathway activated by macrophages (Supplementary 5). The study, which used databases such as PHANTER, allowed us to globally view the biological processes in which the selected proteins are involved. According to Supplementary 6, proteins are involved in signalling processes, the immune system, regulation, localisation, locomotion, etc. New protein-protein interaction analysis was performed, which observed that the selected biomarkers presented a theorical direct molecular relationships between IL-8 (or CXCL-8), IL-6 and CYFRA21 (KRT19). The data gathered from STRING indicated that KRT19 is involved in the organisation of myofibrils alongside KRT8, while CEA plays a very important role in cell adhesion and intracellular signalling processes (Fig. 4, Supplementary 7 and Supplementary 8). Only CYFRA21 can be considered as a good marker for the OPMD/HCG subgroup, by asking the program to indicate more possible nodes/interactions related to this protein we determined a strong cluster composed of cytoskeleton proteins and Mesothelin; a membrane-binding

protein that plays a very important role in the cell adhesion process (Supplementary 7 and Supplementary 9).

4. Discussion

The hypothesis of this study proposed the idea that there are protein-based biomarkers in saliva whose quantitative expression in OPMD is different to that of healthy control groups and malignant lesions, therefore making it possible for reference points of expression to be established. Although we found studies that analysed 11 different biomarkers, it was only possible to perform a meta-analysis on 8 studies with 4 biomarkers.

CEA is a transmembrane glycoprotein that has multiple functions. In OSCC of the tongue, its overexpression correlates with neutrophil infiltration and both are associated with poor prognosis[44]. In our meta-analysis, CEA presented differences in expression between OSCC and OPMD ($p<0.001$). The value of the expression in the Chinese population[36] and the Italian population[43] was 29.3 (CI95%: 11.6–46.9) and 22.2 (CI95%: 4.1–40.3), respectively, with an OE of 25.85 (CI95%: 13.2–38.4). However, CEA did not present a significant difference in expression when comparing OPMD with HCG in the considered studies. The distribution of men and women in the population in Zheng et al's study [36] was similar, although the number of cases of OSCC was almost 4 times higher than those of OPMD (Table 1), which also included the diagnoses of oral lichen planus, OL and erythema, without specifying the number of cases that corresponded to each diagnosis (risk of bias). The salivary levels of CEA in the OSCC group showed a significant association with the tumour stage ($p=0.018$) and lymph node metastasis ($p=0.035$)[36]. On the other hand, more than 70% of the subjects in Airoldi et al's study [43] were male, identifying that 70.6% of the OPMD

corresponded to OL (n=12). In this study, the CEA was significantly greater in OSCC compared to OPMD ($p<0.025$) and HCG (<0.001).

Historically, CYFRA21 has been researched as a biomarker for non-small cell lung carcinoma and it was subsequently evaluated in other malignant[45] and potentially malignant disorders[46]. It may also be a useful biomarker in the head and neck region[47]. The OE for CYFRA21 was highly significant in OSCC compared to OPMD ($p<0.001$), and the latter group also showed significance when compared to HCG ($p=0.017$). In the considered studies, populations of the same origin were studied[37,39]. The OPMD group exhibited an OE which was 3.6 (CI95%: 0.6-6.6) times greater than HCG, despite this, there was a wide heterogeneity in the estimation of the included studies, for Awasthi et al[37] with 2 (CI95%: 0.2–3.7) and Rajkumar et al[39] with 5 (CI95%: 4.9–5.2). The Indian population, which was studied by Awasthi et al[37] was comprised of more than 88% male individuals and it evaluated a quantity of OPMD that was equivalent to 1/3 of OSCC cases, specifying that the OPMD group was comprised of 5 OL, 2 submucous fibrosis, 1 lichen planus and 1 squamous papilloma. CYFRA21 exhibited a positive association with the tumour size, clinical tumour stage and lymph node status. Furthermore, analysis of the ROC curve showed a sensitivity of 90% and a specificity of 97%, with a minimum cut-off value of 8.7ng/mL for the detection of OSCC. The study by Rajkumar et al [39] included an OPMD group that consisted of 50 cases of OL and 50 cases of oral submucous fibrosis, with males representing 60% of the case and a reported rate of 14% of smokers for both OSCC and OL. The analysis of the ROC curve showed specificity and sensitivity of 75%, with a cut-off limit >10.4ng/mL to differentiate OSCC from OPMD, additionally, it

showed a specificity of 95% and a sensitivity of 83.6%, with a cut-off limit >7.91ng/mL for the detection of cases of OSCC in contrast to HCG.

In studies which were performed to determine the interaction between CEA and CYFRA21, the results demonstrated that both present a strong association (Fig. 4) and that they are directly related to a whole cluster of keratins involved in the cellular cytoskeleton, one of the key factors in oral carcinogenesis[48]. All this was in line with studies in which it has been determined that both proteins can be used as diagnostic markers in other epithelial tumours such as lung cancer[49] and head and neck cancer[50], therefore concurring that these two proteins are differentially expressed in healthy and potentially malignant tissues.

The IL-6 salivary mRNA was previously identified as a promising biomarker for OSCC[51], however, studies into the role of protein expression in saliva or OPMD are limited. There was no significant difference in the IL-6 cytokine salivary OE between the OPMD group and HCG. The anticipated means of expression for the studies included in the meta-analysis were considerably different, the highest being 397.8 (CI95%: 381.3–414.3) in the study by Sharma et al[41], compared to 69.4 (CI95%: 56.1–82.6) in the study by Rhodus et al[42], which resulted in a value of extreme heterogeneity. In the study of Rhodus et al, no clinical diagnoses were indicated in the OPMD group, only different histological degrees of dysplasia [42]. The concentration of IL-6 was significantly higher in OSCC and OPMD when compared to HCG ($p<0.001$). On the other hand, in Sharma's study, which was conducted in India, means of salivary expression were almost 10 times higher than those observed in the US population, also were included patients with periodontitis simultaneously to OL, which could be impact in the amount of IL-6 measured. The OL group showed a significantly

higher mean IL-6 compared to the HCG ($p<0.001$). The levels of IL-6 in patients who smoked and had a higher degree of dysplasia were always higher.

Changes in the expression of pro-inflammatory cytokines may play a role in the malignant transformation in OSCC and are likely to be used as diagnostic markers in the serum of patients, due to their excessive production by tumour cells[14], and likewise, they could also act as possible pharmacological targets. IL-8 plays an important role in cancer invasion, angiogenesis and metastasis[52]. In this meta-analysis, we found no significant difference for IL-8 between OSCC and OPMD, nor between OPMD and HCG. Likewise, we did not find any relationship between these pro-inflammatory cytokines and the previously described markers (CEA and CYFRA21), however, as can be seen in Fig. 4, the pro-inflammatory cytokines would form a strong signalling cluster by themselves, through the functional analysis, which has been performed with the DAVID program, which corresponds to the immune response and inflammation signalling pathway (Supplementary 4). The studies included in the meta-analysis showed a very high value of heterogeneity in all subgroups (Fig. 2C and 3C). The study by Gleber-Netto et al[38] from a population in Taiwan was almost exclusively male and had a percentage of smokers which exceeded 80% in all groups. The corresponding clinical diagnoses were not reported in the OPMD group, only those lesions with different degrees of dysplasia were included. On the other hand, the study conducted by Punyani et al[40] in India reported that the OPMD group was comprised of 13 oral submucous fibrosis and 12 OL and that 5 of these had mild dysplasia. The identification of IL-8 was significantly higher in OSCC compared to OPMD ($p<0.0001$) and HCG ($p<0.0001$). The OPMD group presented higher levels of IL-8 compared to HCG, however, this difference was not significant

($p=0.738$). In the study by Rhodus et al[42], the clinical diagnoses were not determined in the OPMD group, only the existence of lesions with different degrees of dysplasia was mentioned. The concentration of IL-8 was significantly higher in OSCC compared to HCG ($p<0.001$), and it was also significantly higher in OPMD compared to HCG ($p<0.05$), although the statistical significance between OSCC and OPMD was not noted.

The main constraints of this study included the lack of clinical-pathological information on the lesions that are mentioned as dysplastic in the OPMD group, and likewise, the degree of dysplasia is not characterised. Additionally, there were considerable discrepancies between the number of analysed cases of OL and OPMD, compared with the cases of OSCC. Very few studies have analysed pure OL groups, independent of other OPMD, therefore making it difficult to draw specific conclusions. Another aspect that must be taken into consideration is that the gender and age of the individuals in the meta-analysis was quite heterogeneous, without there being a strictly matched selection in terms of gender or similar age intervals between all of their study groups. Moreover, not all of the studies reported the presence of risk factors such as smoking (a key factor in oral carcinogenesis) and others included habits that are typical of the geographical region under study, such as chewing betel nut, therefore meaning that extrapolation to a global level was not possible. In terms of the strengths of this study, it is important to highlight the clear and strict inclusion criteria that made it possible to limit analysis biases. Furthermore, even though there were different techniques such as proteomic techniques that allowed for massive quantitative analyses in the literature, this systematic review only included studies in which the quantitative values were representative, reproducible and comparable.

5. Conclusion

This systematic review and meta-analysis provide a summary of the quantitative expression of biomarkers based on salivary proteins. Basing on the results, we can conclude that is quantitively possible to discriminate the OPMD compared to malignant lesions and healthy controls. The CEA and CYFRA21 markers proved to be very useful when differentiating OSCC from OPMD, however, the usefulness of exclusively differentiating the group of leukoplakia from OSCC does not seem totally clear. Besides, CYFRA21 marker proved to be the only protein capable of distinguishing between OPMD and healthy controls. The definition of these protein-based salivary markers as a molecular signature, which identifies potentially malignant lesions, would allow for new, less invasive protocols to be established for early diagnosis and screening, as well as making it possible to determine therapeutic targets that would improve the prognosis of patients with potentially malignant lesions and those at risk.

Sources of support in the form of grants:

Esteban Arroyo received funding for this research from Programa Estudantes-Convênio de Pós-Graduação – PEC-PG, CAPES/CNPq, Brazil (88881.154490/2017-01); Andreia Bufalino received funding for this research from São Paulo Research Foundation (FAPESP) (grant #2017/01438-0) and the National Council for Scientific and Technological Development CNPq (423945/2016-5) ; Darcy Fernandes received funding for this research from São Paulo Research Foundation (FAPESP) (grant #2017/01798-7). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES).

Author contributions

Esteban Arroyo: Conception, Interpretation or analysis of data, Preparation of the manuscript, Revision for important intellectual content.

Sergio Piñeiro Donís: Conception, Interpretation or analysis of data, Preparation of the manuscript, Revision for important intellectual content.

Cintia Chamorro: Interpretation or analysis of data, Preparation of the manuscript.

Monica Oliveira Alves: Interpretation or analysis of data, Preparation of the manuscript.

Xabier Marichalar Mendia: Interpretation or analysis of data, Preparation of the manuscript.

Darcy Fernandes: Interpretation or analysis of data.

Alejandro Lorenzo Pouso: Interpretation or analysis of data, Preparation of the manuscript, Supervision.

Andréia Bufalino: Conception, Supervision.

Susana Bravo López: Conception, Preparation of the manuscript, Supervision.

Pilar Gandara Vila: Conception, Supervision.

Mario Pérez Sayáns: Conception, Interpretation or analysis of data, Preparation of the manuscript, Revision for important intellectual content, Supervision.

REFERENCES

[1] Scully C, Bagan J. Oral squamous cell carcinoma overview. Oral Oncol 2009;45:301–8. https://doi.org/10.1016/j.oraloncology.2009.01.004.

[2] El-Naggar AK, Chan JK, Grandis JR, Takata T, Slootweg PJ, editors. Oral potentially malignant disorders and oral epithelial dysplasia. WHO Classif. Head Neck Tumours. Fourth edition, Lyon, France: International Agency for Research on Cancer; 2017, p. 112–5.

[3] Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer 2019;144:1941–53. https://doi.org/10.1002/ijc.31937.

[4] Sklenicka S, Gardiner S, Dierks EJ, Potter BE, Bell RB. Survival Analysis and Risk Factors for Recurrence in Oral Squamous Cell Carcinoma: Does Surgical Salvage Affect Outcome? J Oral Maxillofac Surg 2010;68:1270–5. https://doi.org/10.1016/j.joms.2009.11.016.

[5] Van der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. Oral Oncol 2009;45:317–23. https://doi.org/10.1016/j.oraloncology.2008.05.016.

[6] Van Der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; present concepts of management. Oral Oncol 2010;46:423–5. https://doi.org/10.1016/j.oraloncology.2010.02.016.

[7] Petti S. Pooled estimate of world leukoplakia prevalence: a systematic review. Oral Oncol 2003;39:770–80. https://doi.org/10.1016/S1368-8375(03)00102-7.

[8] Neville BW, Damm DD, Allen CM, Chi AC. Oral Maxillofac. Pathol., St. Louis, MO: Saunders/Elsevier; 2016, p. 355–63.

[9] Kujan O, Oliver RJ, Khattab A, Roberts SA, Thakker N, Sloan P. Evaluation of a new binary system of grading oral epithelial dysplasia for prediction of malignant transformation. Oral Oncol 2006;42:987–93. https://doi.org/10.1016/j.oraloncology.2005.12.014.

[10] Fleskens SAJHM, Bergshoeff VE, Voogd AC, van Velthuysen M-LF, Bot FJ, Speel E-JM, et al. Interobserver variability of laryngeal mucosal premalignant lesions: a histopathological evaluation. Mod Pathol 2011;24:892–8. https://doi.org/10.1038/modpathol.2011.50.

[11] Sperandio M, Brown AL, Lock C, Morgan PR, Coupland VH, Madden PB, et al. Predictive Value of Dysplasia Grading and DNA Ploidy in Malignant Transformation of Oral Potentially Malignant Disorders. Cancer Prev Res (Phila Pa) 2013;6:822–31. https://doi.org/10.1158/1940-6207.CAPR-13-0001.

[12] Kaufman E, Lamster IB. T HE D IAGNOSTIC A PPLICATIONS OF S ALIVA — A R EVIEW. Crit Rev Oral Biol Med 2002;13:197–212. https://doi.org/10.1177/154411130201300209.

[13] Gröschl M. Saliva: a reliable sample matrix in bioanalytics. Bioanalysis 2017;9:655–68. https://doi.org/10.4155/bio-2017-0010.

[14] Aro K, Wei F, Wong DT, Tu M. Saliva Liquid Biopsy for Point-of-Care Applications. Front Public Health 2017;5. https://doi.org/10.3389/fpubh.2017.00077.

[15] Lorenzo-Pouso AI, Pérez-Sayáns M, Bravo SB, López-Jornet P, García-Vence M, Alonso-Sampedro M, et al. Protein-Based Salivary Profiles as Novel Biomarkers

for Oral Diseases. Dis Markers 2018;2018:1–22. https://doi.org/10.1155/2018/6141845.

[16] Amado F, Lobo MJC, Domingues P, Duarte JA, Vitorino R. Salivary peptidomics. Expert Rev Proteomics 2010;7:709–21. https://doi.org/10.1586/epr.10.48.

[17] Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. Nat Rev Clin Oncol 2017;14:531–48. https://doi.org/10.1038/nrclinonc.2017.14.

[18] Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods 2016;13:731–40. https://doi.org/10.1038/nmeth.3901.

[19] Tyanova S, Cox J. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. In: von Stechow L, editor. Cancer Syst. Biol., vol. 1711, New York, NY: Springer New York; 2018, p. 133–48. https://doi.org/10.1007/978-1-4939-7493-1_7.

[20] Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res 2018;46:D649–55. https://doi.org/10.1093/nar/gkx1132.

[21] Kuhn M, Szklarczyk D, Pletscher-Frankild S, Blicher TH, von Mering C, Jensen LJ, et al. STITCH 4: integration of protein–chemical interactions with user data. Nucleic Acids Res 2014;42:D401–7. https://doi.org/10.1093/nar/gkt1207.

[22] Szklarczyk D, Santos A, von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. Nucleic Acids Res 2016;44:D380–4. https://doi.org/10.1093/nar/gkv1277.

[23] Muetze T, Goenawan IH, Wiencko HL, Bernal-Llinares M, Bryan K, Lynn DJ. Contextual Hub Analysis Tool (CHAT): A Cytoscape app for identifying contextually relevant hubs in biological networks 2016.

[24] Gurevich VV. Protein multi-functionality: introduction. Cell Mol Life Sci 2019;76:4405–6. https://doi.org/10.1007/s00018-019-03271-6.

[25] Higgins JP, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 2011.

[26] Liberati A, Altman D, Tetzlaff J, Mulrow C, Gøtzsche P, Ioannidis J, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. J Clin Epidemiol 2009;62:e1-34. https://doi.org/10.1016/j.jclinepi.2009.06.006.

[27] Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;349:g7647–g7647. https://doi.org/10.1136/bmj.g7647.

[28] Axell T, Pindborg JJ, Smith CJ, Waal I, an International Collaborative Group on Oral White Lesions*. Oral white lesions with special reference to precancerous and tobacco-related lesions: conclusions of an international symposium held in Uppsala, Sweden, May 18-21 1994. J Oral Pathol Med 1996;25:49–54. https://doi.org/10.1111/j.1600-0714.1996.tb00191.x.

[29] Van der Waal I. Historical perspective and nomenclature of potentially malignant or potentially premalignant oral epithelial lesions with emphasis on leukoplakia—some suggestions for modifications. Oral Surg Oral Med Oral Pathol Oral Radiol 2018;125:577–81. https://doi.org/10.1016/j.oooo.2017.11.023.

[30] Leeflang MM, Deeks JJ, Takwoingi Y, Macaskill P. Cochrane diagnostic test accuracy reviews. Syst Rev 2013;2:82. https://doi.org/10.1186/2046-4053-2-82.

[31] Ciapponi A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica 2015:5.

[32] Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, et al. Protocol Update for Large-scale genome and gene function analysis with PANTHER Classification System (v.14.0). Nat Protoc 2019;14:703–21. https://doi.org/10.1038/s41596-019-0128-8.

[33] 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;45:D362–8. https://doi.org/10.1093/nar/gkw937.

[34] Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4:P3.

[35] Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–58. https://doi.org/10.1002/sim.1186.

[36] Zheng J, Sun L, Yuan W, Xu J, Yu X, Wang F, et al. Clinical value of Naa10p and CEA levels in saliva and serum for diagnosis of oral squamous cell carcinoma. J Oral Pathol Med 2018;47:830–5. https://doi.org/10.1111/jop.12767.

[37] Awasthi N. Role of salivary biomarkers in early detection of oral squamous cell carcinoma. Indian J Pathol Microbiol 2017;60:464. https://doi.org/10.4103/IJPM.IJPM_140_16.

[38] Gleber-Netto FO, Yakob M, Li F, Feng Z, Dai J, Kao H-K, et al. Salivary Biomarkers for Detection of Oral Squamous Cell Carcinoma in a Taiwanese Population. Clin Cancer Res 2016;22:3340–7. https://doi.org/10.1158/1078-0432.CCR-15-1761.

[39] Rajkumar K, Ramya R, Nandhini G, Rajashree P, Ramesh Kumar A, Nirmala Anandan S. Salivary and serum level of CYFRA 21-1 in oral precancer and oral squamous cell carcinoma. Oral Dis 2015;21:90–6. https://doi.org/10.1111/odi.12216.

[40] Punyani SR, Sathawane RS. Salivary level of interleukin-8 in oral precancer and oral squamous cell carcinoma. Clin Oral Investig 2013;17:517–24. https://doi.org/10.1007/s00784-012-0723-3.

[41] 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;15:705–14. https://doi.org/10.1007/s00784-010-0435-5.

[42] 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;29:42–5. https://doi.org/10.1016/j.cdp.2004.10.003.

[43] Airoldi M, Piantino P, Pacchioni D, Mastromatteo V, Pedani F, Gandolfo S. Gastrointestinal cancer-associated antigen (GICA) in oral carcinoma. Oral Surg Oral Med Oral Pathol 1986;61:263–7. https://doi.org/10.1016/0030-4220(86)90372-5.

[44] Wang N, Wang Q, Chi J, Xiang F, Lin M, Wang W, et al. Carcinoembryonic antigen cell adhesion molecule 1 inhibits the antitumor effect of neutrophils in tongue squamous cell carcinoma. Cancer Sci 2019;110:519–29. https://doi.org/10.1111/cas.13909.

[45] Molina R, Agusti C, Filella X, Jo J, Joseph J, Giménez N, et al. Study of a new tumor marker, CYFRA 21-1, in malignant and nonmalignant diseases. Tumour Biol J Int Soc Oncodevelopmental Biol Med 1994;15:318–25. https://doi.org/10.1159/000217908.

[46] Camisasca DR, da Rós Gonçalves L, Soares MR, Sandim V, Nogueira FCS, Garcia CHS, et al. A proteomic approach to compare saliva from individuals with and without oral leukoplakia. J Proteomics 2017;151:43–52. https://doi.org/10.1016/j.jprot.2016.07.029.

[47] AlAli AM, Walsh T, Maranzano M. CYFRA 21-1 and MMP-9 as salivary biomarkers for the detection of oral squamous cell carcinoma: a systematic review of diagnostic test accuracy. Int J Oral Maxillofac Surg 2020:S0901502720300369. https://doi.org/10.1016/j.ijom.2020.01.020.

[48] Pitiyage G, Tilakaratne WM, Tavassoli M, Warnakulasuriya S. Molecular markers in oral epithelial dysplasia: review. J Oral Pathol Med 2009;38:737–52. https://doi.org/10.1111/j.1600-0714.2009.00804.x.

[49] Okamura K, Takayama K, Izumi M, Harada T, Furuyama K, Nakanishi Y. Diagnostic value of CEA and CYFRA 21-1 tumor markers in primary lung cancer. Lung Cancer 2013;80:45–9. https://doi.org/10.1016/j.lungcan.2013.01.002.

[50] Barak V, Meirovitz A, Leibovici V, Rachmut J, Peretz T, Eliashar R, et al. The Diagnostic and Prognostic Value of Tumor Markers (CEA, SCC, CYFRA 21-1, TPS) in Head and Neck Cancer Patients. Anticancer Res 2015;35:5519–24.

[51] Márton IJ, Horváth J, Lábiscsák P, Márkus B, Dezső B, Szabó A, et al. Salivary IL-6 mRNA is a Robust Biomarker in Oral Squamous Cell Carcinoma. J Clin Med 2019;8:1958. https://doi.org/10.3390/jcm8111958.

[52] Ralli M, Grasso M, Gilardi A, Ceccanti M, Messina MP, Tirassa P, et al. The role of cytokines in head and neck squamous cell carcinoma: A review. Clin Ter 2020;171:e268–74.

Table 1. Descriptive data from the studies included in the meta-analysis

First Author	Year	Country	Study Subgroup	Proteomic Biomarker	N	Age	Male (%)	Smokers (%)	Mean Expression	Method	Dysplasia
Zheng	2018	China	OSCC	CEA	112	63.2	55.35	-	146.02	ELISA ng/ml	-
			OPMD	CEA	30	61.43	53.33	-	116.72	ELISA ng/ml	-
			HCG	CEA	60	60.57	61.66	-	77.34	ELISA ng/ml	-
Airoldi	1986	Italy	OPMD	CEA	17	52	82.4	82	58.52	ELISA ng/mL	-
			OSCC	CEA	11	66	81.8	81	80.72	ELISA ng/mL	-
			HCG	CEA	19	58	70	42	46.31	ELISA ng/mL	-
Awasthi	2017	India	OSCC	CYFRA 21	30	49.6	93.3	-	17.5	ELISA ng/ml	-
			OPMD	CYFRA 21	9	34.2	88.9	-	5.9	ELISA ng/ml	-
			HCG	CYFRA 21	25	48.1	88	-	3.9	ELISA ng/ml	-
Rajkumar	2015	India	OSCC	CYFRA 21	100	55.5	68	14	17.46	ELISA ng/ml	-
			OPMD	CYFRA 21	100	55.5	58	14	8.15	ELISA ng/ml	-
			HCG	CYFRA 21	100	43	65	0	3.06	ELISA ng/ml	-
Rhodus	2005	USA	OSCC	IL-6	13	59.46	76.92	61.53	0.0882	ELISA pg/ml	-
			OPMD	IL-6	13	56.76	30.76	53.84	0.0708	ELISA pg/ml	Moderate (38.46%) Severe (61.54)

			HCG	IL-6	13	58.7	53.84	0	0.0014	ELISA pg/ml	-
Sharma	2011	India	OL	IL-6	20	50	100	35	0.414	ELISA pg/ml	Mild (40%) Moderate (35%) Severe (25%)
			HCG	IL-6	20	35	100	0	0.0172	ELISA pg/ml	-
Rhodus	2005	USA	OSCC	IL-8	13	59.46	76.92	61.53	3154	ELISA pg/ml	-
			OPMD	IL-8	13	56.76	30.76	53.84	1918	ELISA pg/ml	Moderate (38.46%) Severe (61.54)
			HCG	IL-8	13	58.7	53.84	0	1580	ELISA pg/ml	-
Gleber-Netto	2016	Taiwan	OSCC	IL8	60	51.4	95	83.3	283.75	ELISA pg/ml	-
			OPMD	IL8	60	50.8	95	96.7	140.35	ELISA pg/ml	-
			HCG	IL8	60	50.5	91.7	81.7	127.79	ELISA pg/ml	-
Punyani	2012	India	OSCC	IL8	25	53.2	64	-	1718.610	ELISA pg/mL	-
			OPMD	IL8	25	32.16	76	-	299.513	ELISA pg/mL	Mild (20 %)
			HCG	IL8	25	45.12	76	-	210.096	ELISA pg/mL	-

Note: CEA: carcinoembryonic antigen, CYFRA 21: soluble fragment of cytokeratin 19, IL-6: interleukin-6, IL-8: interleukin-8.

Table 2. Overall expression values

Biomarker	**Study Subgroup**	**OE**	**CI 95% (p)**	**Heterogeneity (I^2, p)**
CEA	OSCC/OPMD	25.854	13.215–38.492 (**p<0.001**) *	Null (I^2=0, p=0.582)
	OPMD/HCG	25.394	-1.221–52.008 (p=0.061)	High (I^2=79.91, p=0.026)
CYFRA21	OSCC/OPMD	9.317	9.014–9.619 (**p<0.001**) *	Null (I^2=0, p=0.428)
	OPMD/HCG	3.679	0.663–6.696 (**p=0.017**) *	High (I^2=91.24, p<0.001)
IL-6	OPMD/HCG	233.561	-88.265–555.387 (p=0.155)	Extreme (I^2=99.89, p<0.001)
IL-8	OSCC/OPMD	912.622	-108.248–1933.492 (p=0.08)	Extreme (I^2=97.69, p<0.001
	OPMD/HCG	45.967	-24.473–116.407 (p=0.201)	Low (I^2=39.30, p=0.193)

*Note: Overall expression values (OE) with a 95%, confidence interval (CI95%) and the heterogeneity (I^2) for each marker in the different analysis subgroups. * Stands for statistically significant differences*

Figures and Supplementary material captions

Fig. 1. Flow chart describing the management of data since the search, data processing and selection of articles for meta-analysis.

Fig. 2. Forest plot with the weighted mean differences, overall expression (OE), 95% confidence interval and heterogeneity for each protein biomarker in saliva for the OSCC/OPMD subgroup analysis. A) CEA. B) CYFRA21. C) IL8.

Fig. 3. Forest plot with the weighted mean differences, overall expression (OE), 95% confidence interval and heterogeneity for each protein biomarker in saliva for the OPMD/HCG subgroup analysis A) CEA. B) CYFRA21. C) IL8. D) IL6

Fig. 4. Protein interaction network of the 4 protein markers selected using the STRING databases.

Supplementary 1. PRISMA checklist

Supplementary 2. Search strategy

Supplementary 3. Analysis of the methodological quality and the risk of bias using the QUADAS-2 scale.

Supplementary 4. DAVID Functional analysis: interactions between the IL system and TNF-α.

Supplementary 5. DAVID Functional analysis: cells and molecules involved in the local acute inflammatory response.

Supplementary 6. Pathway enrichment analysis using PANTHER for the 4 proteins of interest.

Supplementary 7. Protein interaction network by adding more nodes to the human CEA and CYFRA21 proteins determined using STRING databases. MSLN: mesothelin

Supplementary 8. Human CYFRA21 and CEA Protein interaction network using STRING databases.

Supplementary 9. Protein interaction network of CYFRA21 with the other cytoskeleton proteins determined using STRING databases.

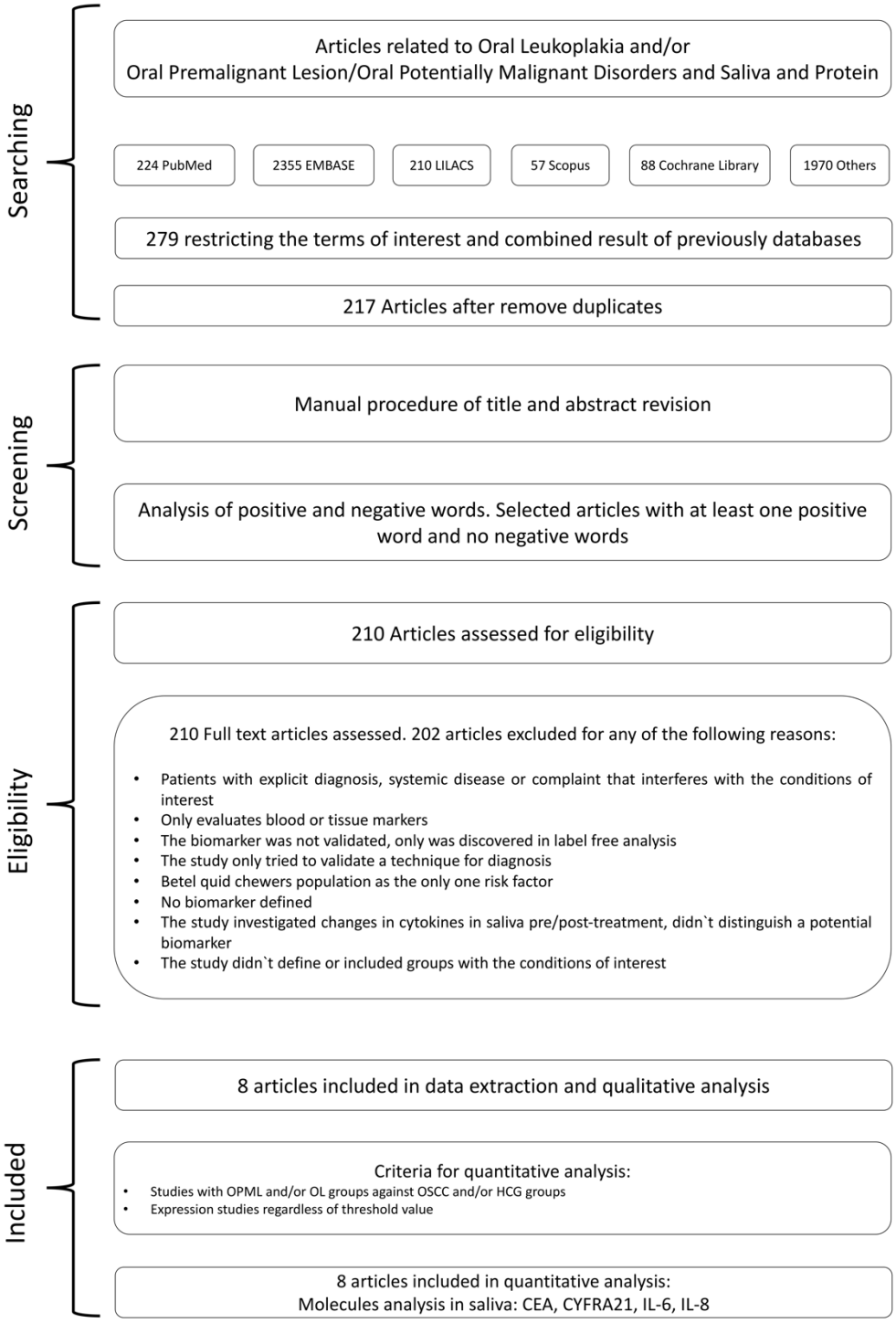


Fig. 1 Flowchart

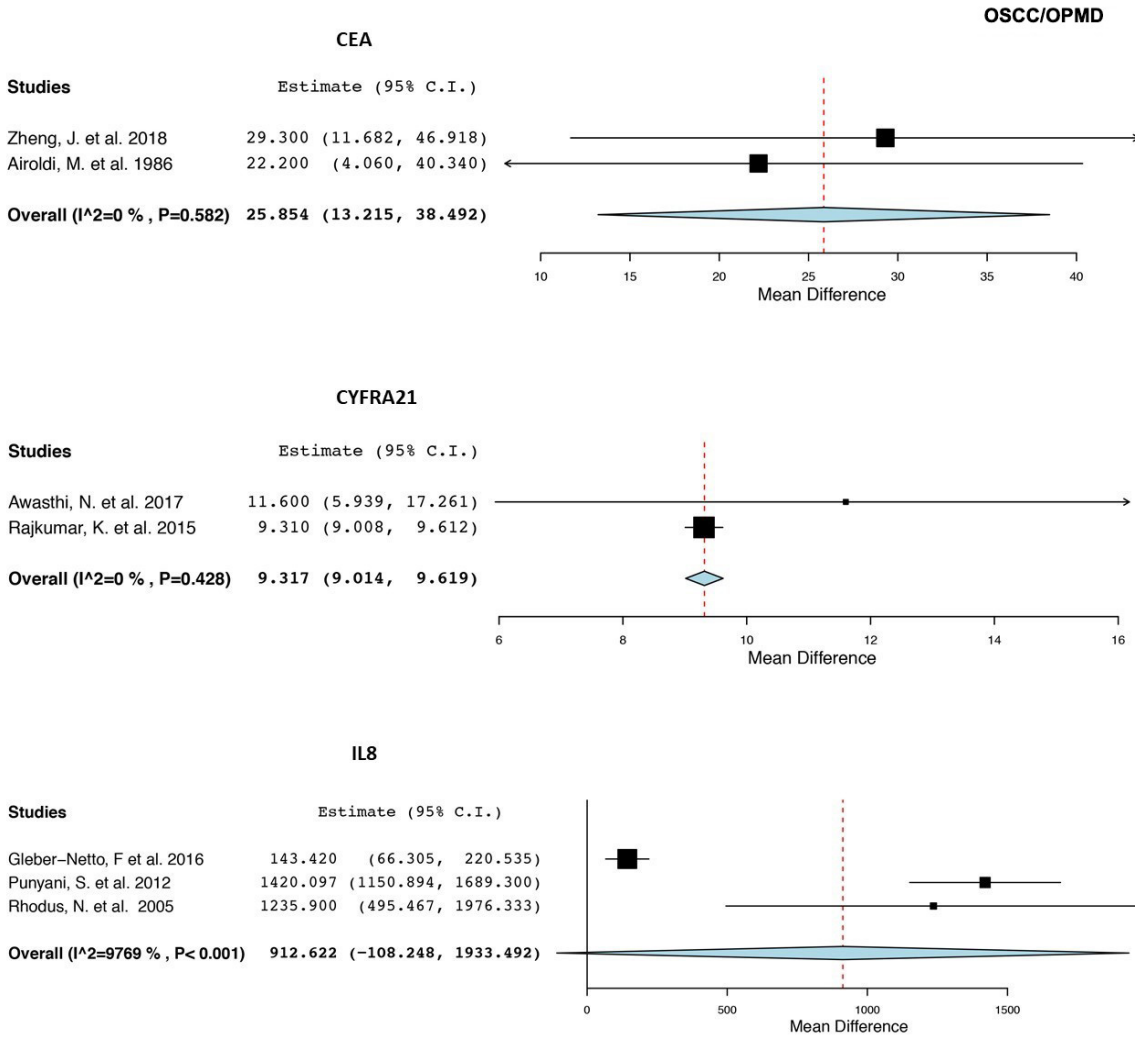


Fig. 2 Forest plot OSCC/OPMD

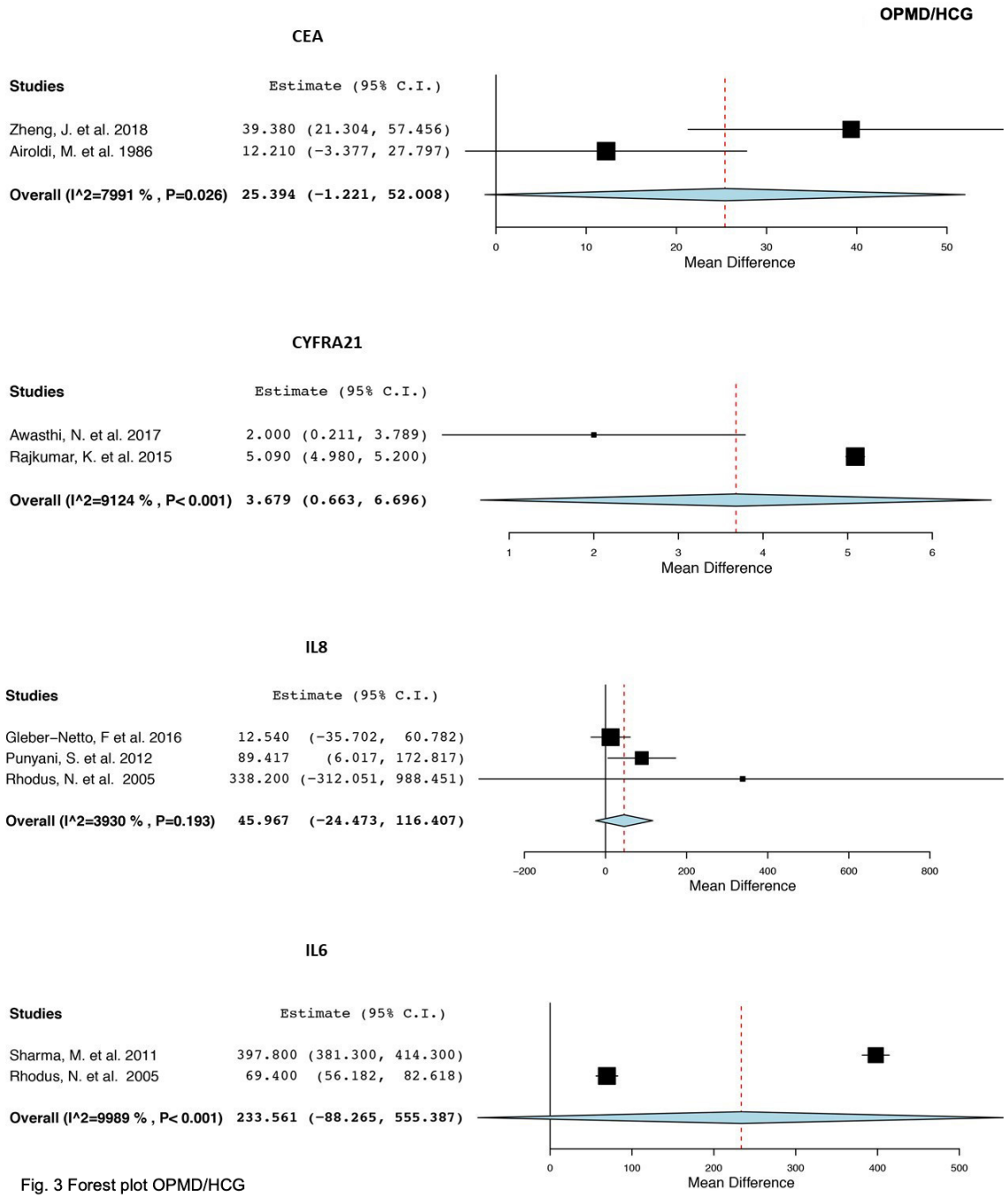


Fig. 3 Forest plot OPMD/HCG

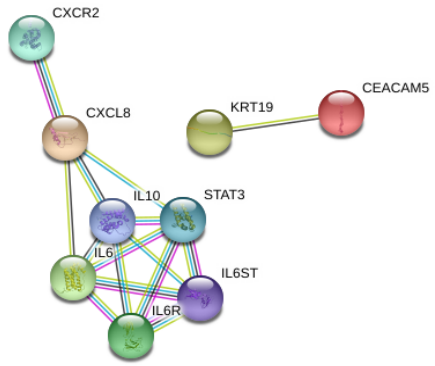


Fig. 4 Protein network of protein markers selected

Supplementary 1. PRISMA checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5-6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	8
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	8 Supplementary 2
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	9
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	9-10 Figure 1
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5-7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis	10 Supplementary 3
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	10-11
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	10-12

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	10 Supplementary 3
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	11-12
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	12
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	12-13 Table 1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	13 Supplementary 3
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	13 Figure 2(A, B, C) Figure 3 (A, B, C, D)
Synthesis of results	21	Present the main results of the review. If meta-analyses are done, include for each, confidence intervals and measures of consistency.	13-14 Table 2
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Supplementary 3
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	14-15 Figure 4 Supplementary 4 Supplementary 5 Supplementary 6 Supplementary 7 Supplementary 8 Supplementary 9
DISCUSSION			
Summary of evidence	n	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	15-16
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	19
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	20
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	20

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

Supplementary 2.1 Search strategy applied to the different databases.

Terms for the target condition		
1. Oral Leukoplakia 2. Oral Premalignant Lesion 3. Oral Potentially Malignant Disorder 4. 1 or 2 or 3 AND		
Terms for the type of oral sample analysed		
5. Saliva AND/OR Protein		
Terms for the index tests		
6. 8-OHdG	49. Esterase	96. Metabolite
7. Activating factor	50. Fibronectin	97. MMP
8. Adipocytokine	51. Gingipain	98. Microglobulin
9. Adiponectin	52. Glucuronidase	99. MIP-1α
10. Albumin	53. Glycosaminoglycan	100.MMP-3
11. Amino acid	54. Glycosidase	101.MMP-7
12. Aminopeptidase	55. Growth factor	102.MMP-8
13. Aminotransferase	56. ERBB2	103.MMP-9
14. Amylase	57. HGF	104.MUC4
15. Antibody	58. HSP27	105.MPO
16. Antitrypsin	59. Hydroxyproline	106.Naa10p
17. Arginase	60. ICTP	107.Neopterin
18. Arylsulfatase	61. IFN-α	108.Neurokinin
19. ASC	62. IFN-γ	109.Nitrate
20. Ascorbate	63. IgA	110.NO
21. BGLAP	64. IgA2	111.Nitrite
22. C4d	65. IgE	112.NLRP3
23. Ca^{+2}	66. IgG	113.OPG
24. Calgranulin	67. IgM	114.OPN
25. Calprotectin	68. IL-1α	115.PDGF-BB
26. CASP1	69. IL-1b	116.Peptidase
27. Cathepsin	70. IL-6	117.Peptide
28. CD14	71. IL-8	118.Peroxidase
29. CEA	72. IL-10	119.PGE2
30. Chemerin	73. IL-17	120.Phosphatase
31. Chemokine	74. IL-18	121.Plasminogen
32. Chitinase	75. IL-21	122.PCT
33. Chondroitin	76. IL-22	123.Progranulin
34. Collagenase	77. IL1RA	124.Prostaglandin
35. Complement C	78. Immunoglobulin	125.Protease
36. Cortisol	79. Interferon	126.Protein
37. Creatine	80. Keratin	127.Proteinase
38. Creatinine	81. L-plastin	128.Pyridinoline
39. CRP	82. LDH	129.RANKL
40. CYFRA21	83. Lactoferrin	130.RANTES
41. Cystatin	84. Laminin	131.Resistin
42. Cytokine	85. LDH	132.S100A8
43. Dehydrogenase	86. Leptin	133.S100A9

44. Dipeptidyl peptidase	87. Leukotriene	134.sCD14
45. EGF	88. Lysozyme	135.SOD
46. Elastase	89. Macroglobulin	136.SPARC
47. EDN1	90. MAF	137.Substance
48. Enzyme	91. Marker	138.TIMP-1
	92. MCP-1	139.TLR-2
	93. MDA	140.TLR-4
	94. Mediator	141.TNF-α
	95. Melatonin	142.VEGF
		143.Visfatin
		144.α-amylase

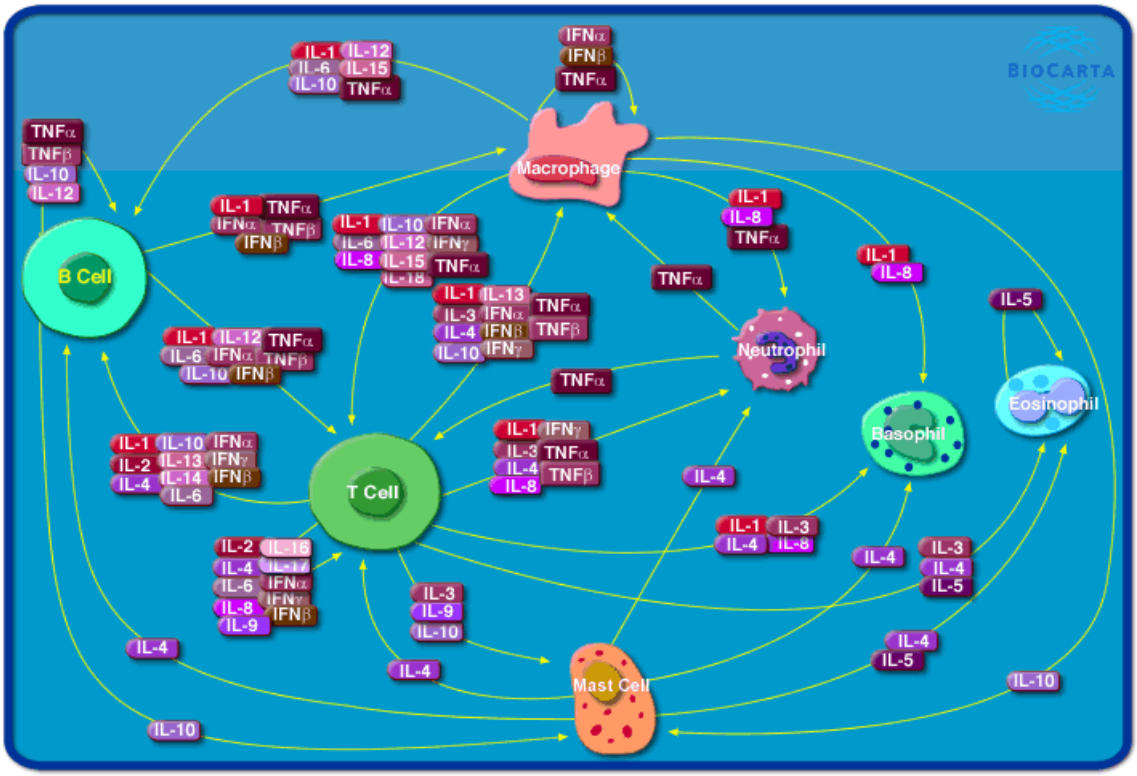
Supplementary 2.2. List of positive and negative words that were used when selecting the candidate articles.

Positive words	Accuracy, sensitivity, specificity, threshold, area under curve, receiver operating, operating characteristic, positive predictive value, negative predictive value, true positive, true negative, false positive, false negative, point of care, chairside test, diagnostic test, prognostic test, logistic regression, canonic correlation, odd ratio, neuronal network, support vector machine, performance measure, predictive model, accurate, prediction, regression, discriminant, cluster, clustering, variance.
Negative words	dog, cat, animal, mouse, rat, vitro, monkey, pig, rabbit, review, meta-analysis, overview.

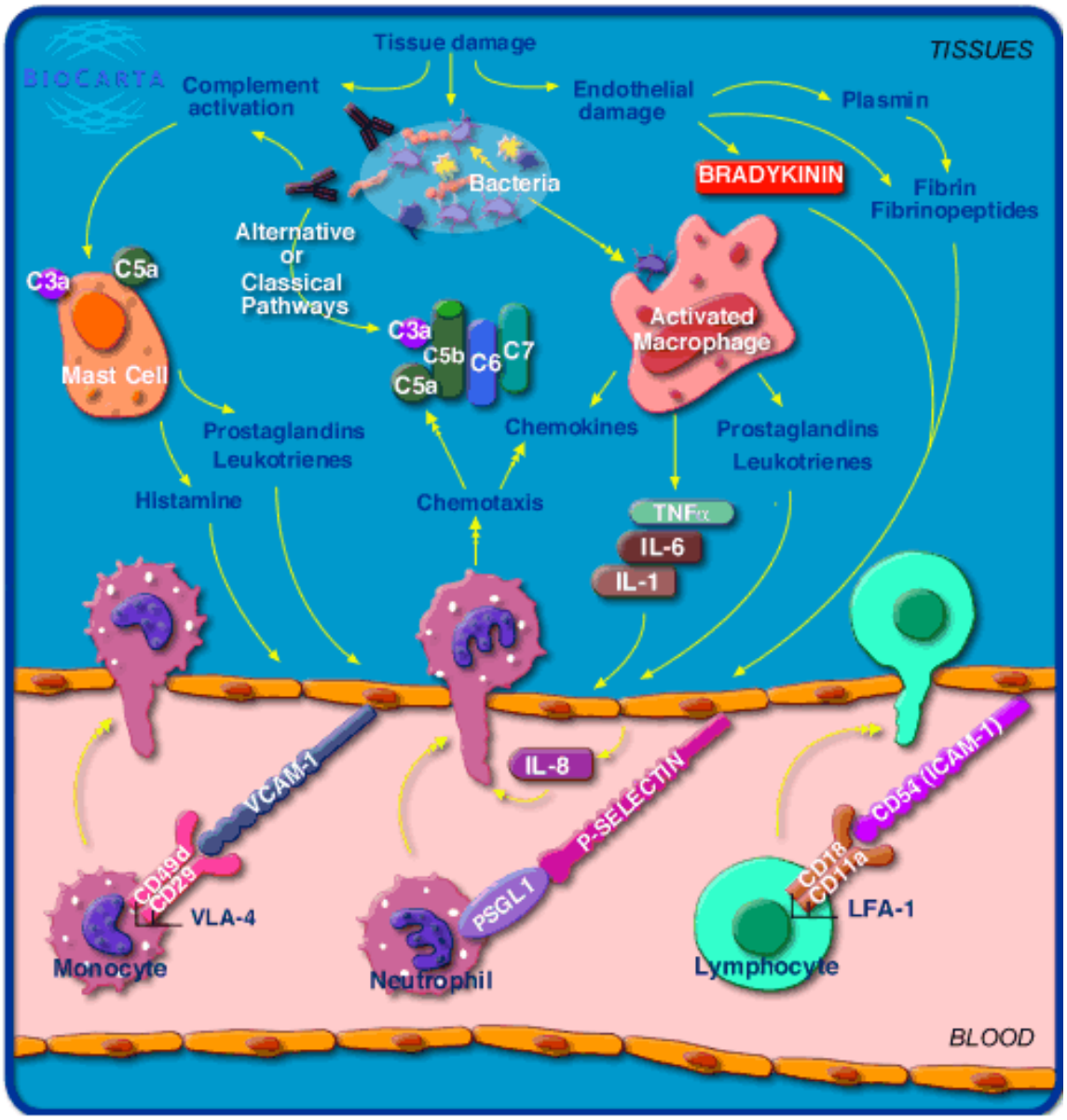
Supplementary 3. Analysis of the methodological quality and the risk of bias using the QUADAS-2 scale

		RISK OF BIAS				CONCERNS REGARDING APPLICABILITY		
STUDY	YEAR	PATIENT SELECTION	INDEX TEST	REFERENCE STANDART	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDART
Zheng, J. *et al.*	2018	HIGH	?	LOW	LOW	LOW	LOW	?
Aiello, M. *et al.*	1986	HIGH	?	LOW	LOW	LOW	LOW	LOW
Awasthi, N. *et al.*	2017	HIGH	?	LOW	LOW	LOW	LOW	LOW
Rajkumar, K. *et al.*	2015	HIGH	?	LOW	LOW	LOW	LOW	LOW
Rhodus, N. L. *et al.*	2005	HIGH	?	LOW	LOW	LOW	LOW	LOW
Sharma, M. *et al.*	2011	HIGH	?	LOW	LOW	LOW	LOW	LOW
Gleber-Netto, F. O. *et al.*	2016	HIGH	?	LOW	LOW	LOW	LOW	LOW
Punyani, S. R. *et al.*	2012	HIGH	?	LOW	LOW	LOW	LOW	LOW

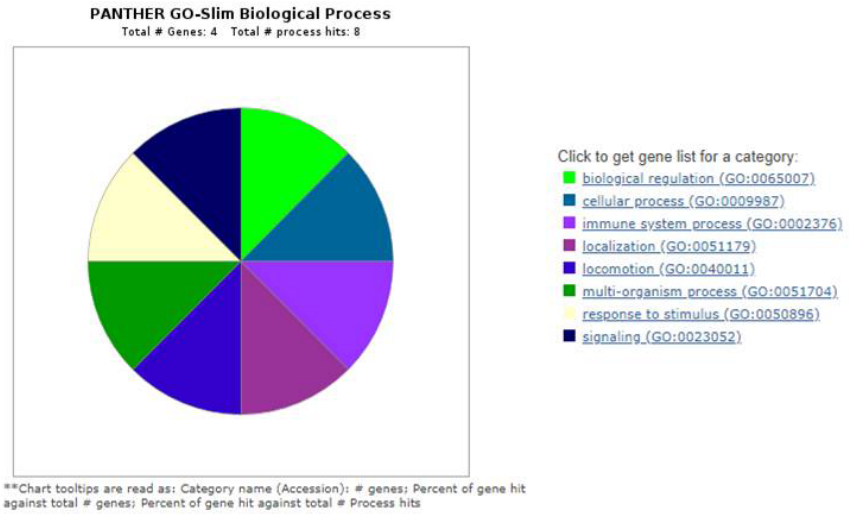
	RISK OF BIAS			CONCERNS REGARDING APPLICABILITY		
	LOW PROBABILITY	HIGH PROBABILITY	UNCERTAN PROBABILITY	LOW PROBABILITY	HIGH PROBABILITY	UNCERTAN PROBABILITY
PATIENT SELECTION	0	8	0	8	0	0
INDEX TEST	0	0	8	8	0	0
REFERENCE STANDART	8	0	0	7	0	1
FLOW AND TIMING	8	0	0	-	-	-



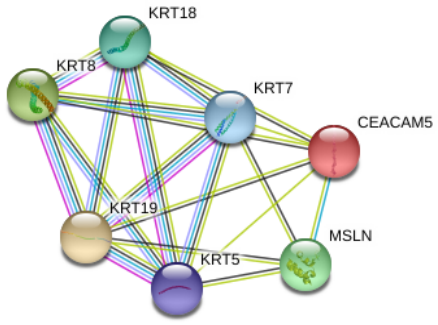
Supplementary 4. DAVID Functional analysis: interactions between the IL system and TNF-α.



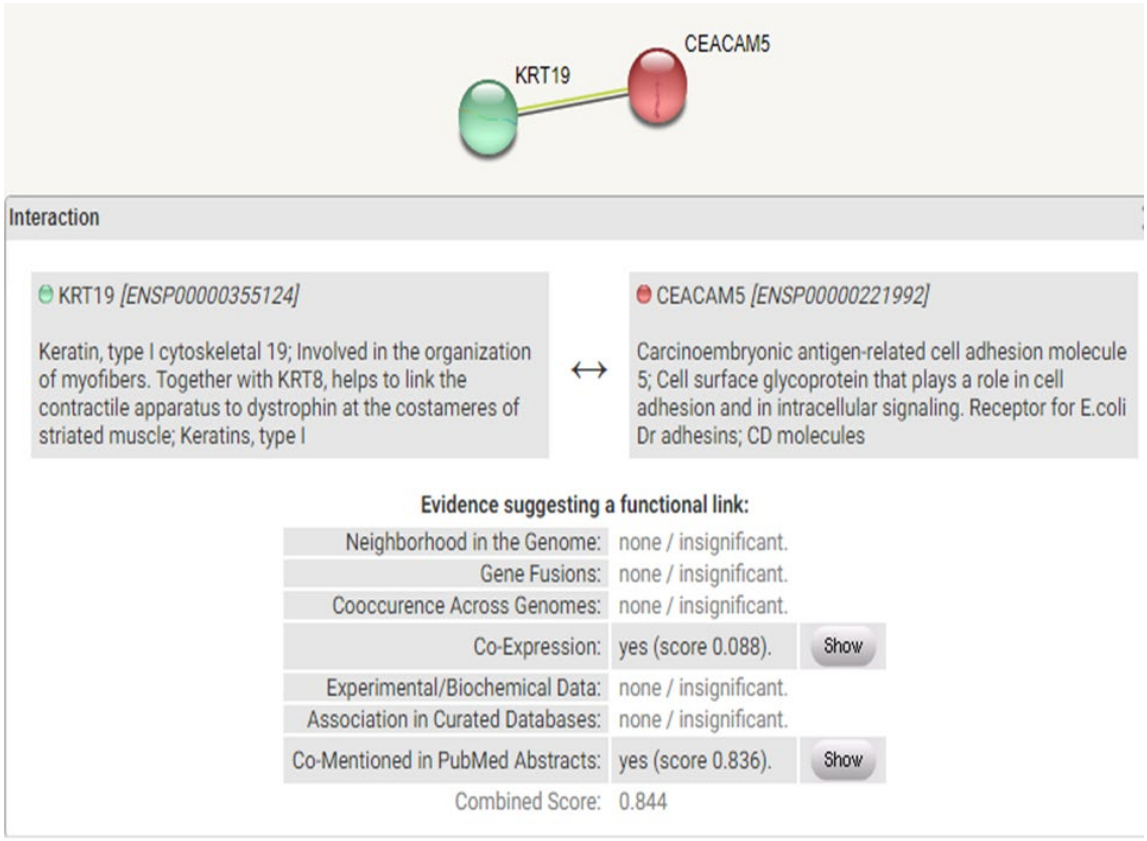
Supplementary 5. DAVID Functional analysis: cells and molecules



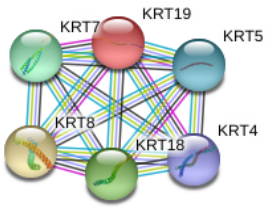
Supplementary 6. Pathway enrichment analysis using PANTHER for the 4 proteins of interest.



Supplementary 7. Protein interaction network of CEA and CYFRA21 proteins



Supplementary 8. CYFRA21 and CEA Protein interaction.



Supplementary 9. Protein interaction network of CYFRA21 with the other cytoskeleton proteins.

3.2 Publicação 2*

TITLE

CD4 & CD20 densities discern Proliferative Verrucous Leukoplakia which mimics Oral Lichen Planus

RUNNING TITLE

CD4 & CD20 densities characterise PVL

AUTHORS

Esteban Arroyo[1], Mariana Paravani Palaçon[1], Camila de Oliveira Barbeiro[1], Túlio Morandin Ferrisse[2], Pilar Gándara Vila[3,4], Mario Pérez Sayáns[3,4], Heitor Albergoni da Silveira[1], Jorge Esquiche León[15], **Andreia Bufalino[1]***.

[1] Department of Diagnosis and Surgery, Araraquara, School of Dentistry, São Paulo State University (Unesp), Araraquara, SP, Brazil.

[2] Department of Dental Materials and Prosthodontics, São Paulo State University (Unesp), School of Dentistry, Araraquara, Brazil

[3] Oral Medicine, Oral Surgery and Implantology Unit (MedOralRes). Faculty of Medicine and Dentistry, Universidade de Santiago de Compostela.

[4] ORALRES GROUP, Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain

[5]Oral Pathology, Department of Stomatology, Public Oral Health, and Forensic Dentistry, Ribeirão Preto Dental School (FORP/USP), University of São Paulo, Ribeirão Preto, SP 14040-904, Brazil.

*** CORRESPONDENCE**

Prof. Dr Andreia Bufalino PhD, MSc, DDS. Address: Department of diagnostic and Surgery, Sao Paulo State University. 1680 Humaita St, 2nd floor, Araraquara, Sao Paulo, Brazil. TLF: (16) 3301 6382. E-mail: andreia.bufalino@unesp.br

Keywords: Proliferative verrucous leukoplakia, Oral Lichen Planus, Immunohistochemistry, Adaptive immunity.

Date of submission: 11th of November

* O artigo segue as normas do periódico Oral Diseases (ISSNs: 1601-0825) ao qual foi submetido.

ABSTRACT

This study aims to compare the inflammatory infiltrate among proliferative verrucous leukoplakia, oral lichenoid lesion and oral lichen planus. 27 proliferative verrucous leukoplakia, 14 oral lichen planus, 14 oral lichenoid lesions and 12 controls were processed for immunohistochemistry of CD3, CD4, CD8 and CD20 expression. Groups and layers (intraepithelial and subepithelial) were statistically compared. The correlation was calculated for CD4+ and CD20+ for each group. SPSS version 20.0 was used with a significance of 95%. Immunohistochemistry analysis for T and B lymphocytes revealed a significant result for CD3+($p=0.001$), CD4+($p<0.0001$) and CD20+($p<0.029$). In the subepithelial layer, CD3+ was significantly higher in oral lichen planus($p<0.001$) and proliferative verrucous leukoplakia($p<0.027$). CD4+ was significantly higher in the oral lichen planus than in the other groups($p<0.05$). CD20+ in the intraepithelial layer had a lower density in proliferative verrucous leukoplakia compared with the control($p<0.0001$). In the subepithelial layer were lower densities of CD20+ in proliferative verrucous leukoplakia($p<0.032$) and oral lichenoid lesion($p<0.010$). It was found a significant correlation for proliferative verrucous leukoplakia between CD4+ and CD20+ cells($p=0.0366$). The different densities in lymphocytes could explain the behaviour of oral lichen planus, as well the higher malignant potential of oral lichenoid lesion and proliferative verrucous leukoplakia.

INTRODUCTION

Oral potentially malignant disorders (OPMD) are clinical conditions that can precede oral squamous cell carcinoma (OSCC) (El-Naggar, Chan, Grandis, Takata, & Slootweg, 2017; Van der Waal, 2009; Van Der Waal, 2010), which represent more than 95% of oral neoplasms, with high prevalence and mortality rates, as well a poor clinical outcome who does not exceed 50% to 60% of surviving in five years (Ferlay et al., 2019; Scully & Bagan, 2009; Sklenicka, Gardiner, Dierks, Potter, & Bell, 2010). Thus, prevention and early diagnosis are extremely necessary to improve the prognosis and survival rates for these patients.

Proliferative verrucous leukoplakia (PVL) is the OPMD with the highest malignant transformation rate (Iocca et al., 2020). Clinically, PVL appears as a non-homogenous leukoplakia that generally affects multiple sites, with slow-growing involvement of contiguous or non-contiguous areas, being resistant to all forms of current treatments(Borgna et al., 2017; Capella, Gonçalves, Abrantes, Grando, & Daniel, 2017; Celentano, Glurich, Borgnakke, & Farah, 2020; Torrejon-Moya, Jané-Salas, & López-López, 2020; Villa et al., 2018). Histopathologically, PVL can range from hyperkeratosis with or without dysplasia to verrucous

hyperplasia, verrucous carcinoma or OSCC, whether in situ or infiltrating(Cerero-Lapiedra, Balade-Martinez, Moreno-Lopez, Esparza-Gomez, & Bagan, 2010; Hansen, Olson, & Silverman, 1985).

Otherwise, oral lichen planus (OLP) is a chronic inflammatory disease of unknown aetiology, commonly characterized by the presence of bilateral white reticular lesions, mostly symmetrical, but others forms like erosive and atrophic can be associated (Giuliani et al., 2019; González-Moles et al., 2019; Neville, Damm, Allen, & Chi, 2016; van der Meij, Schepman, & van der Waal, 2003). Histologically is characterized by the presence of a well-defined bandlike zone of cellular infiltration confined to the superficial part of the connective tissue, consisting mainly of lymphocytes, with signs of "liquefaction degeneration" in the basal cell layer and absence of epithelial dysplasia (van der Meij et al., 2003).

Moreover, oral lichenoid lesion (OLL) is a clinical entity used to describe oral lesions with similar aspects (clinically or histopathologically) of oral lichen planus, caused by an identifiable etiological factor in the majority of cases. Dental restorative materials or systemic medications, as well as secondary to another pathological condition are examples of OLL etiological factors (Cheng, Gould, Kurago, Fantasia, & Muller, 2016; Chiang et al., 2018). In addition, OLL was recently classified as OPMD, after an international seminar on nomenclature and classification, convened by the WHO Collaborating Centre for Oral Cancer, since there is enough evidence of the increased risk of oral cancer among the patients diagnosed with OLL (Warnakulasuriya et al., 2021).

OLP and OLL are considered the OPMD's with the lowest malignant transformation rate (Iocca et al., 2020), although this is still controversial, largely attributable to the different criteria employed for diagnosis, which lead to the creation of various terms related to lichen, even defining versions of OLP that may or may not have determined the malignant transformation potential (El-Naggar et al., 2017; Giuliani et al., 2019; González-Moles et al., 2019; Iocca et al., 2020). Moreover, some studies have shown that PVL, especially in the early stages, may have clinical and microscopic aspects that mimic OLP and/or OLL (Garcia-Pola, Llorente-Pendas, Gonzalez-Garcia, & Garcia-Martin, 2016; Gillenwater, Vigneswaran, Fatani, Saintigny, & El-Naggar, 2013, 2014; Lopes, Feio, Santos-Silva, & Vargas, 2015; Müller, 2011, 2017).

According to WHO, OPMD diagnoses should be based on precise clinical and histopathological findings, but these represents a problem because these methods are poorly reproducible, have high inter-observer variability, and present limitations in the presence of OPMD with similar

histological characteristics (Fleskens et al., 2011; Kujan et al., 2006; Sperandio et al., 2013; Upadhyaya et al., 2020).

Therefore, this study proposes a comparative analysis of the lymphocyte infiltrate of PVL, OLP and OLL in the intraepithelial and subepithelial layers, to provide new tools for differential diagnosis and elucidate the role of lymphocyte infiltrate in this OPMD.

METHODS

This retrospective analysis considered the clinical records and tissues samples of patients that constituted 4 different groups. These were compounded by one control group with 12 samples of oral inflammatory fibrous hyperplasia (OIFH), based on the principles adopted in a previously published work (Ferrisse et al., 2021), also, three experimental groups with 27 samples of PVL, 14 samples of OLP and 14 samples of OLL, from the Faculty of Dentistry of Araraquara (FOAr-UNESP), Faculty of Dentistry of São José dos Campos (FOSJC-UNESP) and Federal University of Rio Grande do Norte (UFRN). All patients in experimental groups were followed up for a minimum of 8 years. The samples on the PVL group had a microscopic characteristic that went from hyperkeratosis and/or acanthosis to verrucous hyperplasia, with or without dysplasia. In experimental groups were excluded samples that microscopically showed carcinoma *in situ*, squamous cell carcinoma or verrucous carcinoma. All samples were analysed for morphological description and diagnosis confirmation by two different examiners. The study was approved by the local Research Ethics Committee (CAAE: 34361814.9.0000.5416). All patients provided written informed consent.

Immunohistochemical methods

For immunohistochemical staining, 3 µm thick sections were placed on silane-coated glass slides. Briefly, the sections were deparaffinized and rehydrated in graded ethanol solutions. After antigen retrieval with EDTA/Tris buffer (pH 9.0) in a microwave oven (1380 W; Panasonic, São Paulo, Brazil), endogenous peroxidase activity was blocked with 20% H2O2 using five cycles of 5 min each. Overnight incubation with the primary antibodies for CD3 (Polyclonal; DakoCytomation, Glostrup, Denmark; 1:500), CD4 (4B12; Leica Biosystems Newcastle, UK; 1:200), CD8 (C8/144B; DakoCytomation, Glostrup, Denmark; 1:400) and CD20 (L26; DakoCytomation, Glostrup, Denmark; 1:2000) diluted in bovine serum albumin (BSA) was followed by incubation with the secondary antibody conjugated with polymer dextran marked with peroxidase (Dako EnVision Labeled Polymer; Dako, Glostrup, Denmark). Reactions were developed with a solution containing 0.6 mg/ml 3,30 -diaminobenzidine tetrahydrochloride (DAB, Sigma, St. Louis, MO, USA) and 0.01% H2O2 and counterstained with Carazzi's

haematoxylin. Thereafter, the sections were dehydrated in a series of graded ethanol solutions, diaphanized and mounted in Canada balsam under cover glasses. Positive and negative controls were included in all reactions.

Immunostaining assessment and statistical analysis

For quantification, photos of five representative fields of the intraepithelial and subepithelial areas were taken separately with Olympus DP25 camera (Olympus, Center Valley, PA, USA) attached to the Nikon Eclipse E600 microscope (Nikon, Tokyo, Japan) using a 20 × objective. Positive cells were counted by two independent examiners, using the Image J software (version 1.52, NIH, Bethesda, MD, USA).

Statistical Analysis

Collected data were tabulated and normality was assessed by the Shapiro-Wilk test and the descriptive analysis of Skewness and Kurtosis. The presence of outliers was checked. The homoscedasticity was checked by the Box's Test of Equality of Covariance Matrices and by Levene's Test of Equality of Error Variance The density of different immunophenotypes of lymphocytes was considered as response variables and the independent factors were grouped (PVL, OLP, OLL and control) and layer (epithelial and subepithelial areas). Thus, multiple analysis of variance (Manova two-way) was applied. In presence of significant results, Tukey post-test was applied. In addition, correlation studies were performed evaluating the density of CD4+ and CD20+ cells for PVL (Pearson's correlation coefficient/r), OLP and OLL (Spearman's correlation coefficient/rs). SPSS version 20.0 was used for statistical analysis with a significance level of 95% (α= 0.05).

RESULTS

Clinical, demographic and histologic characteristics of the study groups

All OLP patients had lesions at multiple sites, buccal mucosa was the most affected site (n=5; 33.3%), followed by tongue (n=4; 27%), gingiva (n=3; 19%) and lips (n=3; 19%). The most prevalent clinical presentation was white reticular lesions associated with the erosive form (n=9; 60%). Every single PVL patient had lesions in multiple sites, with a total lesion size greater than 3 cm. At some point in the clinical course of the disease, 7 patients in the PVL group (46.6%) had microscopic features described as lichenoid in association with leukoplakic lesions (Figure 1). All individuals classified as PVL had lesions that spread and affected new sites in the oral cavity. On average each PVL patient had 2.3 biopsies. All cases in the control group were clinically presented as an irregularly shaped nodule, with a smooth/ulcerated surface and located

in a groove bottom region, directly related to the edge of an improperly adapted removable prosthesis. Table 1 summarizes the clinical and histologic characteristics of the groups.

Assessment of density of lymphocytes.

IHC analysis for T and B lymphocytes revealed an overall reduction of these populations in PVL samples compared to the OLP, OLL and control group (Figure 2). Through the different groups, the T and B cells were seen mainly in the basal and spinous layers of the intraepithelial. While in the subepithelial layer, T cells formed an infiltrate in close relationship with the basal cell layer, whereas B cells were sparser and more located more deeply in the connective tissue.

The homoscedasticity tests indicated a non-homoscedasticity distribution for the groups evaluated ($p<0.05$). Thus, the Pillai's Trace was used for the multivariate test, resulting in a significant interaction between the independent factors ($p<0.0001$). Additionally, it was observed a significant result of density of CD3+ ($p=0.001$), CD4+ ($p<0.000$) and CD20+ cells ($p<0.029$). Tukey post-test for multiple comparisons reveals that in the subepithelial layer the CD3+ cells was significantly higher in OLP ($p<0.001$) and PVL($p<0.027$) compared both with the control group, as well in OLP was higher in comparing to the OLL ($p<0.029$), being that the highest expression of CD3+ cells was in the OLP group. The CD4+ T-cells was significantly higher in the OLP compared with OLL ($p<0.015$), PVL ($p<0.000$) and control ($p<0.001$). The CD20+ cells in the intraepithelial layer had a significant lower quantification in PVL compared with control ($p<0.0001$). In the subepithelial layer were found lower densities of CD20+ cells in PVL compared with OLP ($p<0.032$), as well OLL in comparison with the control group ($p<0.010$) (Figure 3).

In the correlation studies performed for CD4+ and CD20+ cells, was noted significant result for PVL group ($p=0.0366/rs=0.4115$) while for OLP ($p=0.8990/rs= -0.0359$) and OLL ($p=0.1292/r=0.3957$) were not significant results (Figure 4).

DISCUSSION

It has been hypothesized that OPMD probably represents a state of balance between the activation of the immune system and the proliferation of dysplastic cells (Dunn, Old, & Schreiber, 2004a, 2004b; Gomes et al., 2016; Mantovani, Allavena, Sica, & Balkwill, 2008; Schreiber, Old, & Smyth, 2011; Zitvogel, Tesniere, & Kroemer, 2006). The results obtained in this study allow for a better understanding of the role of the immune system through the T and B lymphocytes in the PVL, OLL as well OLP, three OPMD with different malignant transformation rates; 49.5%, 3.8% and 1.4%, respectively (Iocca et al., 2020).

The recognition of tumour antigens includes several entities of the immune system; activated macrophages, natural killer cells, CD4+ and CD8+ T lymphocytes, as well B lymphocytes and also plasma cells specifics for tumour peptides, which were found in dysplastic and neoplastic tissues (Zitvogel et al., 2006; Dunn et al., 2004a, 2004b; Mantovani et al., 2008; Schreiber et al., 2011; He et al., 2014; Mandal et al., n.d.). For different types of carcinomas was described that this type of immune cell infiltrates leads to a better clinical outcome (Lee, Zakka, Mihm, & Schatton, 2016; Mellman, Coukos, & Dranoff, 2011; Nosho et al., 2010; Schreiber et al., 2011; Zhu, Lin, Qiao, Xu, & Zou, 2015; Zitvogel et al., 2006). On the other hand, some researchers point to infiltrations of cytotoxic T-cells are associated with a good prognosis while regulatory T-lymphocytes suppress anti-tumour response and because of that, it is related to the worst prognosis(Zou, 2006). In addition, the accuracy of the role of B-cells in tumours is still unfamiliar and can vary according to the type of tumour(Largeot, Pagano, Gonder, Moussay, & Paggetti, 2019).

There is evidence that suggests oral leukoplakia which transformed into OSCC expressed a significantly higher density of CD3+cells in contrast with oral leukoplakia which did not undergo a malignant process (Chen, Tan, Zhang, He, & Zhou, 2019). Thus, high levels of CD3+ cells rates could be a monitoring parameter for malignant transformation. This argumentation could be linked with the results obtained in the present study, which shows higher quantities of CD3+ cells in the intraepithelial layer for PVL, and significantly higher expression of CD3+ in the subepithelial layer for PVL and OLP. However, this would not explain the behaviour of OLL, who shows conflicts in the literature about the malignant transformation rates seen in it, which is even greater than OLP (El-Naggar et al., 2017; Iocca et al., 2020).

Moreover, the malignancy potential and the participation of the immune system in this process for OLL is also still controversial (Thomson, Goodson, & Smith, 2018). The results of this research show significantly lower quantities of CD3+ and CD4+ in the subepithelial layer than OLP, reflecting the loss of capability to respond against tumoral cells. This fact could be explaining the higher malignant behaviour of OLL compared with OLP. Also, the criteria for OLL diagnosis is not standardized in the literature and this can impact the estimation rates of the malignant potential and follow up periods of this type of OPMD(González-Moles et al., 2019).

Furthermore, it was shown a significantly high number of CD4+ in OSCC with metastases and was related to the action of CD4+ T lymphocytes in favour of the invasion and metastasis by promoting the pro-tumoral function of tumour-associated macrophages (Stasikowska-Kanicka, Wągrowska-Danilewicz, & Danilewicz, 2018). The results of this study showed significantly

lower densities in the subepithelial layer of CD4+ in PVL and OLL groups, this could be related to the fact that those lesions have the malignant potential, but in fact, there are not running through a malignant course.

On the other hand, different researches had characterized the lymphocytic infiltrate in OLP mainly as CD4+ cells, independently of the clinical presentation (Omar, Hietanen, Kero, Lukinmaa, & Hagström, 2009; Zhou, Sugerman, Savage, Walsh, & Seymour, 2002). Moreover, the aetiology of OLP is still not fully understood but, some studies pointed out that T lymphocytes represent the greater number of cells related to the pathogenesis of OLP (Chiang et al., 2018; Sugerman et al., 2002). For instance, there have been described four main mechanisms to explain the aetiology of OLP, which three of them related to the action of lymphocytes; they run through antigen-specific mechanisms include antigen presentation by keratinocytes and Langerhans cells to CD4+ helper and CD8+ cytotoxic T lymphocytes to activate these two types of T cells(Chiang et al., 2018; Chitturi, 2014; Córdova, Rubio, & Echeverría, 2013; Lavanya, Jayanthi, Rao, & Ranganathan, 2011; Mutafchieva, Draganova-Filipova, Zagorchev, & Tomov, 2018; Payeras, Cherubini, Figueiredo, & Salum, 2013). Furthermore, has been shown that low levels of CD8+ cells could be a marker of clinical remission of OLP (Enomoto et al., 2018), which is related to the mechanisms mentioned above. Altogether, seems to agree with the observations of this study, which show a significantly higher expression of CD4+ in the subepithelial layer of OLP compared with the other lesions.

Otherwise, different studies had related lower density of CD8+ in OSCC than dysplastic lesions (Pellicioli et al., 2017) as well, a decreased number of CD8+ in OSCC with poor clinical outcome(Maleki et al., 2011; Stasikowska-Kanicka et al., 2018). Despite the present study does not show significant differences in CD8+ in the intraepithelial and subepithelial layers, was possible to note a discreet lower density of CD8+ T lymphocytes in the PVL and OLL samples compared with OLP and control. Different studies had been shown a significantly lower number of CD8+ cells in OPMD with the malignant process involved in comparison with OPMD which never transformed and OSCC (Savina & Amigorena, 2007). This could be consistent with our data since PVL has the highest and OLL has a higher malignant potential than OLP, and this two OPMD probably are losing the immune capacity to self-defence (Iocca et al., 2020).

At last, there are conflicts in the literature concerning B-lymphocytes in the tumorigenesis process(Hadler-Olsen & Wirsing, 2019; He et al., 2014), some evidence pointed out that B-cells may participate in mediating antitumor immunity (Mandal et al., n.d.; Schmidt et al., 2008) and others describe that B-cells can regulate the T-cells by suppressing their cytotoxic activity and

induce CD4+ cells death (Sarvaria, Madrigal, & Saudemont, 2017; Yarchoan et al., 2020). The CD20+ cells infiltrating in the tumour was correlated with better clinical outcomes in OSCC, and could act by; somatic mutation, clonal expansion, intraclonal variation and isotype switching; reflecting an adaptive immune response in the OSCC microenvironment (Quan et al., 2016; Suárez-Sánchez et al., 2021; Taghavi, Mohsenifar, Baghban, & Arjomandkhah, 2018). Also, there is evidence supporting an increasing expression of B-lymphocytes infiltration in the progression to dysplasia and carcinoma in the oral epithelium (Gannot, Gannot, Vered, Buchner, & Keisari, 2002). The present study observed in the intraepithelial layer an absence of B lymphocytes through CD20+ cells in PVL lesion as well OLP and OLL; this could be related to the loss of the tools to self-defend by the humoral response against the tumoral cells. Nonetheless, there is evidence that the presence of CD20+ cells and plasma cells in the OLP does not show a correlation with all the histological features, only was described a higher intensity of CD20+ cells expression in the pattern of keratosis of OLP (Mahdavi, Aminishakib, & Soltani, 2020).

There is evidence showing in non-small cell lung cancer (NSCLC) better survival rates for lesions with high intensities of stromal CD4+ and CD20+ (Al-Shibli et al., 2008), in this study was noted a significant correlation of CD4+ and CD20+ cells in the PVL group ($p=0.0366/rs=0.4115$), this relation could be reinforcing the idea of self-defence mechanism of PVL to prevent malignant transformation process. In addition, further cohort studies evaluating the prognosis of patients affected by PVL with correlation levels of CD4+ and CD20+ cells are expected to strengthen the idea of this mechanism.

Therefore, the results of this study suggest that immune evasion through the reduction of the T and B lymphocytes seems to contribute to the susceptibility of the malignant transformation process in PVL, because the elimination of tumour cells requires the participation of both innate and adaptive immunity pathways, through the integrated humoral and cellular response (Chaves et al., 2019), which also could explain the difference in malignant transformation rates and the behaviour seen between PVL and OLL as well OLP.

CONCLUSIONS

The OLP shows an immune response pattern mediated mainly by T lymphocytes, which could explain their etiopathology, while the OLL shows a decreasing of T and B lymphocytes nearly of PVL behaviour, and this, in turn, shows a remarkable reduction of both T and B lymphocytes, which can work in favour to the malignant transformation process, supporting the behaviour seen in this OPMD.

ACKNOWLEDGEMENTS

N/A

SOURCES OF SUPPORT IN THE FORM OF GRANTS

Esteban Arroyo received funding for this research from Estudantes-Convênio de Pós-Graduação – PEC-PG, CAPES/CNPq, Brazil (88881.154490/2017-01); Andreia Bufalino received funding for this research from São Paulo Research Foundation (FAPESP) (grant #2017/01438-0 and #2021/01191-0) and the National Council for Scientific and Technological Development CNPq (423945/2016-5). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES)

CONFLICTS OF INTEREST: none to declare

SAMPLE CREDIT AUTHOR STATEMENT:

Esteban Arroyo: Investigation; Visualization; Funding acquisition; Writing - original draft; Writing - review & editing**; Mariana Paravani Palaçon:** Formal analysis; Investigation; Methodology**; Camila de Oliveira Barbeiro:** Formal analysis; Investigation; Methodology**; Túlio Morandin Ferrisse:** Data curation; Formal analysis; Software; Supervision; Validation; Visualization; Writing - review & editing**; Pilar Gándara Vila:** Supervision; Writing - review & editing; **Mario Pérez Sayáns:** Supervision; Visualization; Writing - review & editing**; Andreia Bufalino:** Conceptualization; Funding acquisition; Investigation; Project administration; Resources; Supervision; Writing - review & editing

REFERENCES

Al-Shibli, K. I., Donnem, T., Al-Saad, S., Persson, M., Bremnes, R. M., & Busund, L.-T. (2008). Prognostic Effect of Epithelial and Stromal Lymphocyte Infiltration in Non–Small Cell Lung Cancer. *Clinical Cancer Research*, *14*(16), 5220–5227. doi: 10.1158/1078-0432.CCR-08-0133

Borgna, S. C., Clarke, P. T., Schache, A. G., Lowe, D., Ho, M. W., McCarthy, C. E., … Shaw, R. J. (2017). Management of proliferative verrucous leukoplakia: Justification for a conservative approach. *Head & Neck*, *39*(10), 1997–2003. doi: 10.1002/hed.24845

Capella, D. L., Gonçalves, J. M., Abrantes, A. A. A., Grando, L. J., & Daniel, F. I. (2017). Proliferative verrucous leukoplakia: Diagnosis, management and current advances. *Brazilian Journal of Otorhinolaryngology*, *83*(5), 585–593. doi: 10.1016/j.bjorl.2016.12.005

Celentano, A., Glurich, I., Borgnakke, W. S., & Farah, C. S. (2020). World Workshop on Oral Medicine VII: Prognostic biomarkers in oral leukoplakia and proliferative verrucous leukoplakia—A systematic review of retrospective studies. *Oral Diseases*, odi.13363. doi: 10.1111/odi.13363

Cerero-Lapiedra, R., Balade-Martinez, D., Moreno-Lopez, La., Esparza-Gomez, G., & Bagan, Jv. (2010). Proliferative verrucous leukoplakia: A proposal for diagnostic criteria. *Medicina Oral Patología Oral y Cirugia Bucal*, e839–e845. doi: 10.4317/medoral.15.e839

Chaves, A. L. F., Silva, A. G., Maia, F. M., Lopes, G. F. M., de Paulo, L. F. B., Muniz, L. V., … de Azambuja Ribeiro, R. I. M. (2019). Reduced CD8+ T cells infiltration can be associated to a malignant transformation in potentially malignant oral epithelial lesions. *Clinical Oral Investigations*, *23*(4), 1913–1919. doi: 10.1007/s00784-018-2622-8

Chen, X.-J., Tan, Y.-Q., Zhang, N., He, M.-J., & Zhou, G. (2019). Expression of programmed cell death-ligand 1 in oral squamous cell carcinoma and oral leukoplakia is associated with disease progress and CD8+ tumor-infiltrating lymphocytes. *Pathology - Research and Practice*, *215*(6), 152418. doi: 10.1016/j.prp.2019.04.010

Cheng, Y.-S. L., Gould, A., Kurago, Z., Fantasia, J., & Muller, S. (2016). Diagnosis of oral lichen planus: A position paper of the American Academy of Oral and Maxillofacial Pathology. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, *122*(3), 332–354. doi: 10.1016/j.oooo.2016.05.004

Chiang, C.-P., Yu-Fong Chang, J., Wang, Y.-P., Wu, Y.-H., Lu, S.-Y., & Sun, A. (2018). Oral lichen planus—Differential diagnoses, serum autoantibodies, hematinic deficiencies, and

management. *Journal of the Formosan Medical Association = Taiwan Yi Zhi*, *117*(9), 756–765. doi: 10.1016/j.jfma.2018.01.021

Chitturi, R. T. (2014). Oral Lichen Planus: A Review of Etiopathogenesis, Clinical, Histological and Treatment Aspects. *JBR Journal of Interdisciplinary Medicine and Dental Science*, *02*(05). doi: 10.4172/2376-032X.1000142

Córdova, P., Rubio, A., & Echeverría, P. (2013). Oral lichen planus: A look from diagnosis to treatment. *Journal Oral Of Research*, *3*(1), 62–67. doi: 10.17126/joralres.2014.015

Dunn, G. P., Old, L. J., & Schreiber, R. D. (2004a). The immunobiology of cancer immunosurveillance and immunoediting. *Immunity*, *21*(2), 137–148. doi: 10.1016/j.immuni.2004.07.017

Dunn, G. P., Old, L. J., & Schreiber, R. D. (2004b). The Three Es of Cancer Immunoediting. *Annual Review of Immunology*, *22*(1), 329–360. doi: 10.1146/annurev.immunol.22.012703.104803

El-Naggar, A. K., Chan, J. K., Grandis, J. R., Takata, T., & Slootweg, P. J. (Eds.). (2017). Oral potentially malignant disorders and oral epithelial dysplasia. In *WHO Classification of Head and Neck Tumours* (Fourth edition, pp. 112–115). Lyon, France: International Agency for Research on Cancer.

Enomoto, A., Sato, E., Yasuda, T., Isomura, T., Nagao, T., & Chikazu, D. (2018). Intraepithelial CD8+ lymphocytes as a predictive diagnostic biomarker for the remission of oral lichen planus. *Human Pathology*, *74*, 43–53. doi: 10.1016/j.humpath.2017.12.008

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Piñeros, M., … Bray, F. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *International Journal of Cancer*, *144*(8), 1941–1953. doi: 10.1002/ijc.31937

Ferrisse, T. M., de Oliveira, A. B., Palaçon, M. P., Silva, E. V., Massucato, E. M. S., de Almeida, L. Y., … Bufalino, A. (2021). The role of CD68+ and CD163+ macrophages in

immunopathogenesis of oral lichen planus and oral lichenoid lesions. *Immunobiology*, *226*(3), 152072. doi: 10.1016/j.imbio.2021.152072

Fleskens, S. A. J. H. M., Bergshoeff, V. E., Voogd, A. C., van Velthuysen, M.-L. F., Bot, F. J., Speel, E.-J. M., … Slootweg, P. (2011). Interobserver variability of laryngeal mucosal premalignant lesions: A histopathological evaluation. *Modern Pathology*, *24*(7), 892–898. doi: 10.1038/modpathol.2011.50

Gannot, G., Gannot, I., Vered, H., Buchner, A., & Keisari, Y. (2002). Increase in immune cell infiltration with progression of oral epithelium from hyperkeratosis to dysplasia and carcinoma. *British Journal of Cancer*, *86*(9), 1444–1448. doi: 10.1038/sj.bjc.6600282

Garcia-Pola, M., Llorente-Pendas, S., Gonzalez-Garcia, M., & Garcia-Martin, J. (2016). The development of proliferative verrucous leukoplakia in oral lichen planus. A preliminary study. *Medicina Oral Patología Oral y Cirugia Bucal*, e328–e334. doi: 10.4317/medoral.20832

Gillenwater, A. M., Vigneswaran, N., Fatani, H., Saintigny, P., & El-Naggar, A. K. (2013). Proliferative verrucous leukoplakia (PVL): A review of an elusive pathologic entity! *Advances in Anatomic Pathology*, *20*(6), 416–423. doi: 10.1097/PAP.0b013e3182a92df1

Gillenwater, A. M., Vigneswaran, N., Fatani, H., Saintigny, P., & El-Naggar, A. K. (2014). Proliferative verrucous leukoplakia: Recognition and differentiation from conventional leukoplakia and mimics. *Head & Neck*, *36*(11), 1662–1668. doi: 10.1002/hed.23505

Giuliani, M., Troiano, G., Cordaro, M., Corsalini, M., Gioco, G., Lo Muzio, L., … Lajolo, C. (2019). Rate of malignant transformation of oral lichen planus: A systematic review. *Oral Diseases*, *25*(3), 693–709. doi: 10.1111/odi.12885

Gomes, J. O., de Vasconcelos Carvalho, M., Fonseca, F. P., Gondak, R. O., Lopes, M. A., & Vargas, P. A. (2016). CD1a+ and CD83+ Langerhans cells are reduced in lower lip squamous cell carcinoma. *Journal of Oral Pathology & Medicine: Official Publication of the*

International Association of Oral Pathologists and the American Academy of Oral Pathology, *45*(6), 433–439. doi: 10.1111/jop.12389

González-Moles, M. Á., Ruiz-Ávila, I., González-Ruiz, L., Ayén, Á., Gil-Montoya, J. A., & Ramos-García, P. (2019). Malignant transformation risk of oral lichen planus: A systematic review and comprehensive meta-analysis. *Oral Oncology*, *96*, 121–130. doi: 10.1016/j.oraloncology.2019.07.012

Hadler-Olsen, E., & Wirsing, A. M. (2019). Tissue-infiltrating immune cells as prognostic markers in oral squamous cell carcinoma: A systematic review and meta-analysis. *British Journal of Cancer*, *120*(7), 714–727. doi: 10.1038/s41416-019-0409-6

Hansen, L. S., Olson, J. A., & Silverman, S. (1985). Proliferative verrucous leukoplakia. A long-term study of thirty patients. *Oral Surgery, Oral Medicine, and Oral Pathology*, *60*(3), 285–298. doi: 10.1016/0030-4220(85)90313-5

He, Y., Qian, H., Liu, Y., Duan, L., Li, Y., & Shi, G. (2014). The Roles of Regulatory B Cells in Cancer. *Journal of Immunology Research*, *2014*, 1–7. doi: 10.1155/2014/215471

Iocca, O., Sollecito, T. P., Alawi, F., Weinstein, G. S., Newman, J. G., De Virgilio, A., … Shanti, R. M. (2020). Potentially malignant disorders of the oral cavity and oral dysplasia: A systematic review and meta-analysis of malignant transformation rate by subtype. *Head & Neck*, *42*(3), 539–555. doi: 10.1002/hed.26006

Kujan, O., Oliver, R. J., Khattab, A., Roberts, S. A., Thakker, N., & Sloan, P. (2006). Evaluation of a new binary system of grading oral epithelial dysplasia for prediction of malignant transformation. *Oral Oncology*, *42*(10), 987–993. doi: 10.1016/j.oraloncology.2005.12.014

Largeot, A., Pagano, G., Gonder, S., Moussay, E., & Paggetti, J. (2019). The B-side of Cancer Immunity: The Underrated Tune. *Cells*, *8*(5), E449. doi: 10.3390/cells8050449

Lavanya, N., Jayanthi, P., Rao, U. K., & Ranganathan, K. (2011). Oral lichen planus: An update on pathogenesis and treatment. *Journal of Oral and Maxillofacial Pathology : JOMFP*, *15*(2), 127–132. doi: 10.4103/0973-029X.84474

Lee, N., Zakka, L. R., Mihm, M. C., & Schatton, T. (2016). Tumour-infiltrating lymphocytes in melanoma prognosis and cancer immunotherapy. *Pathology*, *48*(2), 177–187. doi: 10.1016/j.pathol.2015.12.006

Lopes, M. A., Feio, P., Santos-Silva, A. R., & Vargas, P. A. (2015). Proliferative verrucous leukoplakia may initially mimic lichenoid reactions. *World Journal of Clinical Cases : WJCC*, *3*(10), 861–863. doi: 10.12998/wjcc.v3.i10.861

Mahdavi, N., Aminishakib, P., & Soltani, N. (2020). Presence of B cells and plasma cells in Oral Lichen Planus. *Journal of Dentistry*, *21*(3). doi: 10.30476/dentjods.2020.81804.0

Maleki, S., Schlecht, N. F., Keller, C., Diaz, J., Moss, J., Prystowsky, M. B., … Brandwein-Gensler, M. (2011). Lymphocytic Host Response to Oral Squamous Cell Carcinoma: An Adaptive T-Cell Response at the Tumor Interface. *Head and Neck Pathology*, *5*(2), 117–122. doi: 10.1007/s12105-011-0247-1

Mandal, R., Şenbabaoğlu, Y., Desrichard, A., Havel, J. J., Dalin, M. G., Riaz, N., … Morris, L. G. T. (n.d.). The head and neck cancer immune landscape and its immunotherapeutic implications. *JCI Insight*, *1*(17), e89829. doi: 10.1172/jci.insight.89829

Mantovani, A., Allavena, P., Sica, A., & Balkwill, F. (2008). Cancer-related inflammation. *Nature*, *454*(7203), 436–444. doi: 10.1038/nature07205

Mellman, I., Coukos, G., & Dranoff, G. (2011). Cancer immunotherapy comes of age. *Nature*, *480*(7378), 480–489. doi: 10.1038/nature10673

Müller, S. (2011). Oral manifestations of dermatologic disease: A focus on lichenoid lesions. *Head and Neck Pathology*, *5*(1), 36–40. doi: 10.1007/s12105-010-0237-8

Müller, S. (2017). Oral lichenoid lesions: Distinguishing the benign from the deadly. *Modern Pathology: An Official Journal of the United States and Canadian Academy of Pathology, Inc*, *30*(s1), S54–S67. doi: 10.1038/modpathol.2016.121

Mutafchieva, M. Z., Draganova-Filipova, M. N., Zagorchev, P. I., & Tomov, G. T. (2018). Oral Lichen Planus – Known and Unknown: A Review. *Folia Medica*, *60*(4), 528–535. doi: 10.2478/folmed-2018-0017

Neville, B. W., Damm, D. D., Allen, C. M., & Chi, A. C. (2016). In *Oral and maxillofacial pathology* (pp. 355–363). St. Louis, MO: Saunders/Elsevier.

Nosho, K., Baba, Y., Tanaka, N., Shima, K., Hayashi, M., Meyerhardt, J. A., … Ogino, S. (2010). Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: Cohort study and literature review. *The Journal of Pathology*, *222*(4), 350–366. doi: 10.1002/path.2774

Omar, A. A. H., Hietanen, J., Kero, M., Lukinmaa, P.-L., & Hagström, J. (2009). Oral lichen planus and chronic junctional stomatitis: Differences in lymphocyte subpopulations. *Acta Odontologica Scandinavica*, *67*(6), 366–369. doi: 10.1080/00016350903136605

Payeras, M. R., Cherubini, K., Figueiredo, M. A., & Salum, F. G. (2013). Oral lichen planus: Focus on etiopathogenesis. *Archives of Oral Biology*, *58*(9), 1057–1069. doi: 10.1016/j.archoralbio.2013.04.004

Pellicioli, A. C. A., Bingle, L., Farthing, P., Lopes, M. A., Martins, M. D., & Vargas, P. A. (2017). Immunosurveillance profile of oral squamous cell carcinoma and oral epithelial dysplasia through dendritic and T-cell analysis. *Journal of Oral Pathology & Medicine*. doi: 10.1111/jop.12597

Quan, H., Fang, L., Pan, H., Deng, Z., Gao, S., Liu, O., … Tang, Z. (2016). An adaptive immune response driven by mature, antigen-experienced T and B cells within the microenvironment

of oral squamous cell carcinoma. *International Journal of Cancer*, *138*(12), 2952–2962. doi: 10.1002/ijc.30019

Sarvaria, A., Madrigal, J. A., & Saudemont, A. (2017). B cell regulation in cancer and anti-tumor immunity. *Cellular and Molecular Immunology*, *14*(8), 662–674. doi: 10.1038/cmi.2017.35

Savina, A., & Amigorena, S. (2007). Phagocytosis and antigen presentation in dendritic cells. *Immunological Reviews*, *219*(1), 143–156. doi: 10.1111/j.1600-065X.2007.00552.x

Schmidt, M., Böhm, D., Törne, C. von, Steiner, E., Puhl, A., Pilch, H., … Gehrmann, M. (2008). The Humoral Immune System Has a Key Prognostic Impact in Node-Negative Breast Cancer. *Cancer Research*, *68*(13), 5405–5413. doi: 10.1158/0008-5472.CAN-07-5206

Schreiber, R. D., Old, L. J., & Smyth, M. J. (2011). Cancer Immunoediting: Integrating Immunity's Roles in Cancer Suppression and Promotion. *Science*, *331*(6024), 1565–1570. doi: 10.1126/science.1203486

Scully, C., & Bagan, J. (2009). Oral squamous cell carcinoma overview. *Oral Oncology*, *45*(4–5), 301–308. doi: 10.1016/j.oraloncology.2009.01.004

Sklenicka, S., Gardiner, S., Dierks, E. J., Potter, B. E., & Bell, R. B. (2010). Survival Analysis and Risk Factors for Recurrence in Oral Squamous Cell Carcinoma: Does Surgical Salvage Affect Outcome? *Journal of Oral and Maxillofacial Surgery*, *68*(6), 1270–1275. doi: 10.1016/j.joms.2009.11.016

Sperandio, M., Brown, A. L., Lock, C., Morgan, P. R., Coupland, V. H., Madden, P. B., … Odell, E. W. (2013). Predictive Value of Dysplasia Grading and DNA Ploidy in Malignant Transformation of Oral Potentially Malignant Disorders. *Cancer Prevention Research*, *6*(8), 822–831. doi: 10.1158/1940-6207.CAPR-13-0001

Stasikowska-Kanicka, O., Wągrowska-Danilewicz, M., & Danilewicz, M. (2018). T cells are involved in the induction of macrophage phenotypes in oral leukoplakia and squamous cell

carcinoma-a preliminary report. *Journal of Oral Pathology & Medicine*, *47*(2), 136–143. doi: 10.1111/jop.12657

Suárez-Sánchez, F. J., Lequerica-Fernández, P., Rodrigo, J. P., Hermida-Prado, F., Suárez-Canto, J., Rodríguez-Santamarta, T., … de Vicente, J. C. (2021). Tumor-Infiltrating CD20+ B Lymphocytes: Significance and Prognostic Implications in Oral Cancer Microenvironment. *Cancers*, *13*(3), 395. doi: 10.3390/cancers13030395

Sugerman, P. B., Savage, N. W., Walsh, L. J., Zhao, Z. Z., Zhou, X. J., Khan, A., … Bigby, M. (2002). The pathogenesis of oral lichen planus. *Critical Reviews in Oral Biology and Medicine: An Official Publication of the American Association of Oral Biologists*, *13*(4), 350–365. doi: 10.1177/154411130201300405

Taghavi, N., Mohsenifar, Z., Baghban, A. A., & Arjomandkhah, A. (2018). CD20+ Tumor Infiltrating B Lymphocyte in Oral Squamous Cell Carcinoma: Correlation with Clinicopathologic Characteristics and Heat Shock Protein 70 Expression. *Pathology Research International*, *2018*, 1–7. doi: 10.1155/2018/4810751

Thomson, P. J., Goodson, M. L., & Smith, D. R. (2018). Potentially malignant disorders revisited-The lichenoid lesion/proliferative verrucous leukoplakia conundrum. *Journal of Oral Pathology & Medicine*, *47*(6), 557–565. doi: 10.1111/jop.12716

Torrejon-Moya, A., Jané-Salas, E., & López-López, J. (2020). Clinical manifestations of oral proliferative verrucous leukoplakia: A systematic review. *Journal of Oral Pathology & Medicine*, *49*(5), 404–408. doi: 10.1111/jop.12999

Upadhyaya, J. D., Fitzpatrick, S. G., Cohen, D. M., Bilodeau, E. A., Bhattacharyya, I., Lewis, J. S., … Thompson, L. D. R. (2020). Inter-observer Variability in the Diagnosis of Proliferative Verrucous Leukoplakia: Clinical Implications for Oral and Maxillofacial Surgeon Understanding: A Collaborative Pilot Study. *Head and Neck Pathology*, *14*(1), 156–165. doi: 10.1007/s12105-019-01035-z

van der Meij, E. H., Schepman, K.-P., & van der Waal, I. (2003). The possible premalignant character of oral lichen planus and oral lichenoid lesions: A prospective study. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics*, *96*(2), 164–171. doi: 10.1016/s1079-2104(03)00305-6

Van der Waal, I. (2009). Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. *Oral Oncology*, *45*(4–5), 317–323. doi: 10.1016/j.oraloncology.2008.05.016

Van Der Waal, I. (2010). Potentially malignant disorders of the oral and oropharyngeal mucosa; present concepts of management. *Oral Oncology*, *46*(6), 423–425. doi: 10.1016/j.oraloncology.2010.02.016

Villa, A., Menon, R., Kerr, A., De Abreu Alves, F., Guollo, A., Ojeda, D., & Woo, S. (2018). Proliferative leukoplakia: Proposed new clinical diagnostic criteria. *Oral Diseases*, *24*(5), 749–760. doi: 10.1111/odi.12830

Warnakulasuriya, S., Kujan, O., Aguirre-Urizar, J. M., Bagan, J. V., González-Moles, M. Á., Kerr, A. R., … Johnson, N. W. (2021). Oral potentially malignant disorders: A consensus report from an international seminar on nomenclature and classification, convened by the WHO Collaborating Centre for Oral Cancer. *Oral Diseases*, *27*(8), 1862–1880. doi: 10.1111/odi.13704

Yarchoan, M., Ho, W. J., Mohan, A., Shah, Y., Vithayathil, T., Leatherman, J., … Jaffee, E. M. (2020). Effects of B cell–activating factor on tumor immunity. *JCI Insight*, *5*(10). doi: 10.1172/jci.insight.136417

Zhou, X. J., Sugerman, P. B., Savage, N. W., Walsh, L. J., & Seymour, G. J. (2002). Intra-epithelial CD8+ T cells and basement membrane disruption in oral lichen planus: CD8+ cells, MCs and BM breaks in OLP. *Journal of Oral Pathology & Medicine*, *31*(1), 23–27. doi: 10.1046/j.0904-2512.2001.10063.x

Zhu, S., Lin, J., Qiao, G., Xu, Y., & Zou, H. (2015). Differential regulation and function of tumor-infiltrating T cells in different stages of breast cancer patients. *Tumor Biology*, *36*(10), 7907–7913. doi: 10.1007/s13277-015-3507-y

Zitvogel, L., Tesniere, A., & Kroemer, G. (2006). Cancer despite immunosurveillance: Immunoselection and immunosubversion. *Nature Reviews. Immunology*, *6*(10), 715–727. doi: 10.1038/nri1936

Zou, W. (2006). Regulatory T cells, tumour immunity and immunotherapy. *Nature Reviews. Immunology*, *6*(4), 295–307. doi: 10.1038/nri1806

TABLE LEGENDS

Table 1. Clinical and histopathological features of groups

FIGURES LEGENDS

Figure 1. Representative clinical and histological images from PVL patients included in this study. Clinical findings of PVL lesion at the time of diagnoses, mimicking an oral lichen planus **(a)**, and its respective histopathological findings showing acanthosis and subepithelial lichenoid inflammatory infiltrate in basal cell layer, (H&E, 10x magnification) and in detail the liquefaction degeneration (H&E, 20x magnification) **(b)**. The same case after 4 years of follow-up showing a white verrucous plaque**(c)**, and its histopathological findings displaying verrucous hyperplasia and prominent keratosis (H&E, 10x magnification).

Figure 2. Representative images of the immunohistochemical staining pattern for CD3+, CD8+ and CD20+ in all studied groups (20x magnification).

Figure 3. CD3+, CD4+, CD8+ and CD20+ intraepithelial (a, b, c, d) and subepithelial (e, f, g, h) in all studied groups, quantified by 5 representative high-resolution fields (400x); *$p<0.05$.

Figure 4. Correlation analysis of CD4+ and CD20+ density cells for a)OLL, b)OLP and c)PVL; *$p<0.05$.

Group	n	Female %	Mean Age	Dysplasia n (%)
OIFH	12	86.7	44.4	-
OLP	14	66.6	48.0	-
OLL	14	80.0	53.3	
PVL	27	60.0	64.6	26 (74.2%) †

†High grade dysplasia

Table 1. Clinical and histopathological features of groups

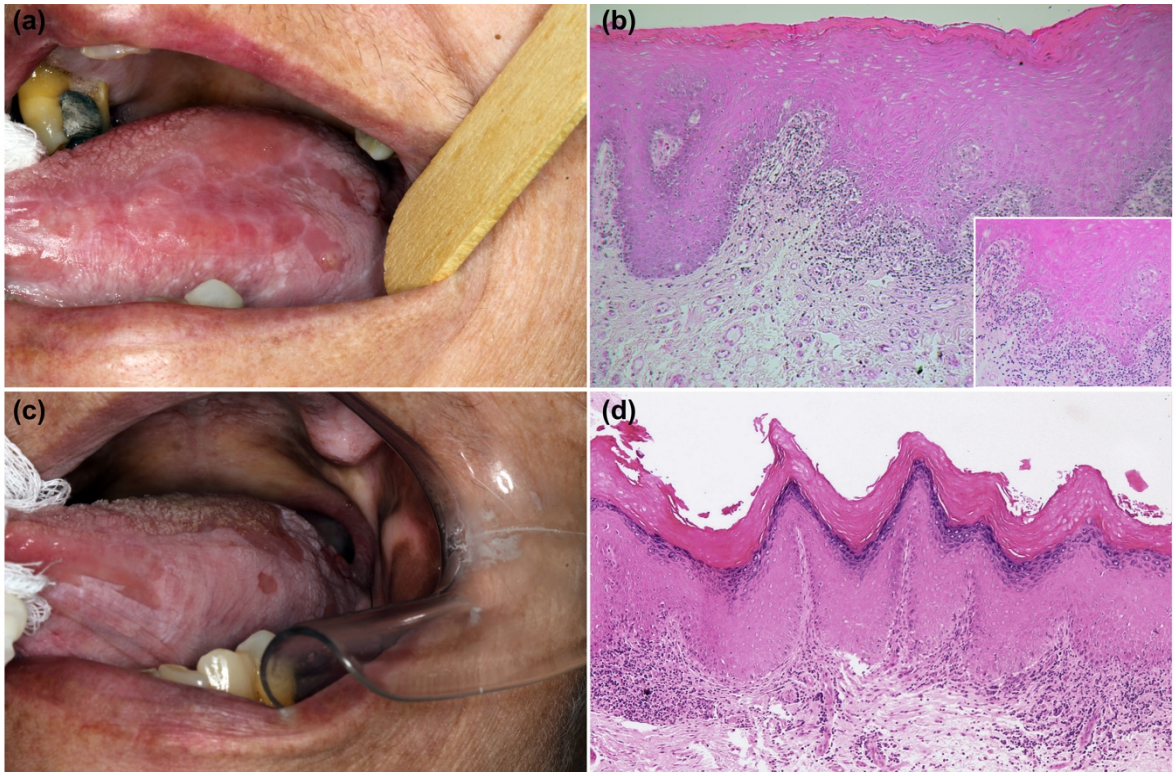


Figure 1. Representative clinical and histological images from PVL patients included in this study. Clinical findings of PVL lesion at the time of diagnoses, mimicking an oral lichen planus **(a)**, and its respective histopathological findings showing acanthosis and subepithelial lichenoid inflammatory infiltrate in basal cell layer, (H&E, 10x magnification) and in detail the liquefaction degeneration (H&E, 20x magnification) **(b)**. The same case after 4 years of follow-up showing a white verrucous plaque**(c)**, and its histopathological findings displaying verrucous hyperplasia and prominent keratosis (H&E, 10x magnification).

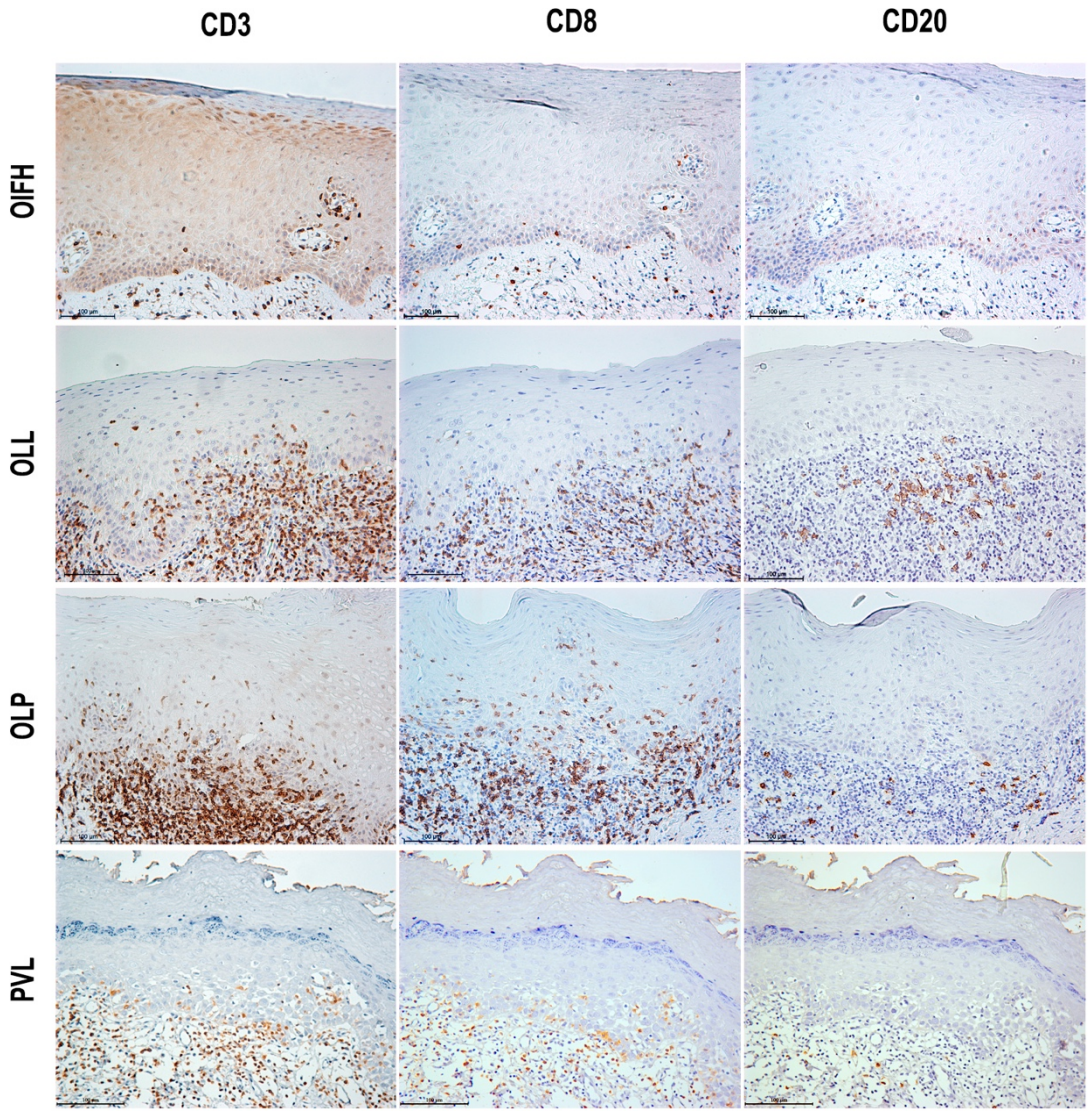


Figure 2. Representative images of the immunohistochemical staining pattern for CD3+, CD8+ and CD20+ in all studied groups (20x magnification).

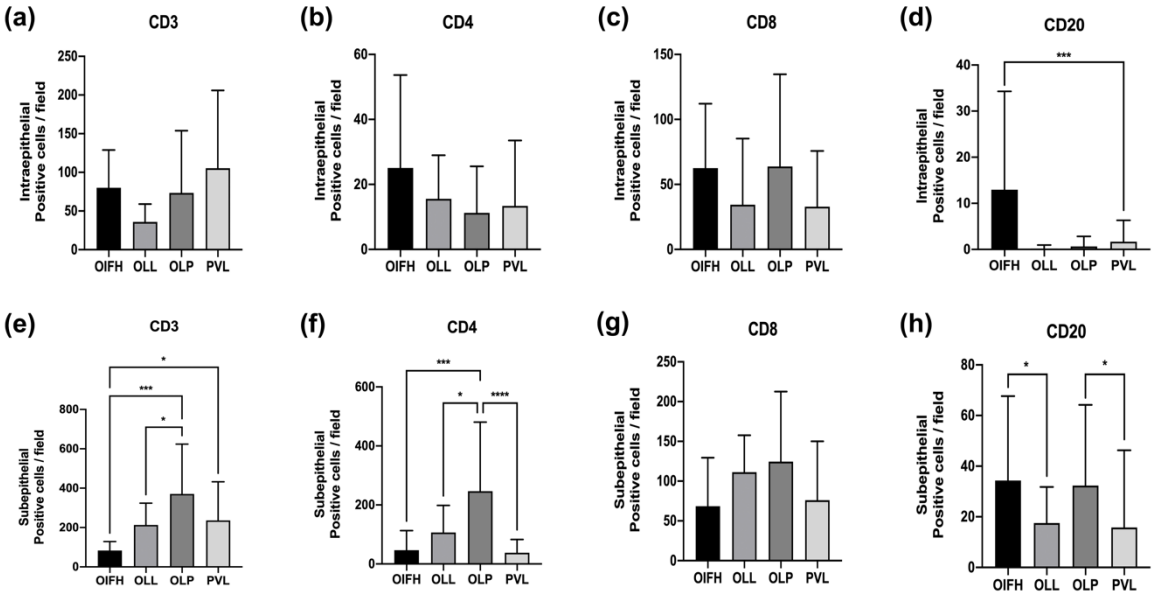


Figure 3. CD3+, CD4+, CD8+ and CD20+ intraepithelial (a, b, c, d) and subepithelial (e, f, g, h) in all studied groups, quantified by 5 representative high-resolution fields (400x); *p<0.05.

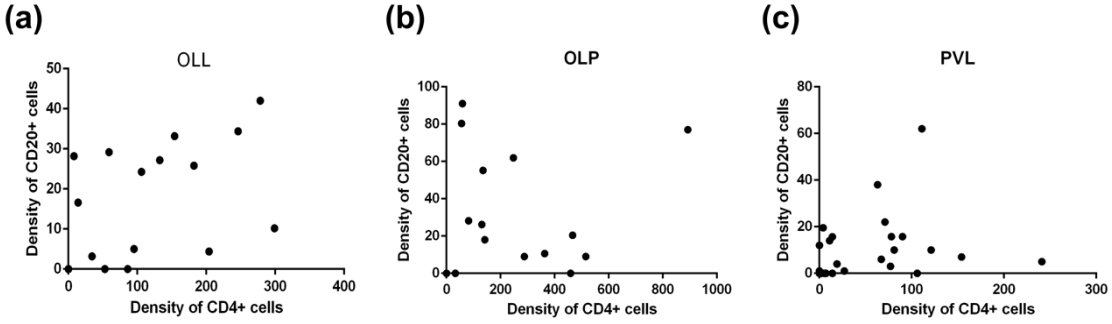


Figure 4. Correlation analysis of CD4+ and CD20+ density cells for a)OLL, b)OLP and c)PVL; *p<0.05.

3.3 Publicação 3*

TITLE

Proteomic analysis on tissue and saliva by LC-MS/MS reveals CALR, YWHAQ and GNB2L1 as putative biomarkers for early diagnosis and therapy intervention in Proliferative Verrucous Leukoplakia

SHORT RUNNING TITLE: Mass spectrometry-based proteomics reveals biomarkers for Proliferative Verrucous Leukoplakia

AUTHORS

Esteban Alexis Arroyo[1], Mario Pérez Sayáns[2,3*], Susana Bravo López[4], Camila de Oliveira Barbeiro[1], Mariana Paravani Palaçon[1], Darcy Fernandes[1], Cintia M Chamorro Petronacci, María García Vence[4], María del Pilar Chantada Vázquez[4], Andrés Blanco Carrión[2], José M Suárez Peñaranda[5], Abel García García[2], Pilar Gándara Vila[2,3], Janete Días Almeida[6], Giovani Carlo Veríssimo da Costa[7], Fábio César Sousa Nogueira[7], Joseph Albert Medeiros Evaristo[7], Denise de Abreu Pereira[9], Mirjami Rintala[10], Tuula Salo[11,12,13], Jaana Rautava[10,11], Elena Padín Iruegas[14], Monica G Oliveira Alves[15], Túlio Morandin Ferrisse[16], Heitor Albergoni da Silveira[1], Jorge Esquiche León[17], Evânio Vilela Silva[1], Luciana Yamamoto Almeida[1], **Andreia Bufalino[1]**.

[1] Department of Diagnosis and Surgery, Araraquara, School of Dentistry, São Paulo State University (Unesp), Araraquara, SP, Brazil.

[2] Oral Medicine, Oral Surgery and Implantology Unit (MedOralRes). Faculty of Medicine and Dentistry, University of Santiago de Compostela.

[3] Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain.

* Este artigo ainda está sobre revisão previa a submissão do mesmo.

[4] Proteomic Unit. Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain.

[5] Servicio de Anatomia Patológica, Hospital Clinico Universitario de Santiago, Choupana s/n Santiago de Compostela.

[6] Department of Bioscience and Buccal Diagnosis, São José dos Campos, Science and Technologies Institute, São Paulo State University (Unesp), São José dos Campos, SP, Brazil.

[7] Department of Basic Sciences, Nova Friburgo Health Institute, Univ. Federal Fluminense, Nova Friburgo, Rio de Janeiro, Brazil.

[8] Laboratory of Proteomics, LADETEC, Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil.

[9] Program on Cellular and Molecular Oncobiology, Research Coordination, National Institute of Cancer/INCA, Rio de Janeiro, Brazil.

[10] Department of Oral Pathology, University of Turku, Turku, Finland.

[11] Department of Oral and Maxillofacial Diseases, Clinicum, University of Helsinki, and Department of Pathology, HUSLAB, Helsinki, Finland.

[12] Department of Cancer and Translational Research Unit, University of Oulu, Oulu, Finland.

[13] Medical Research Center Oulu, Oulu, Finland.

[14]Human Anatomy and Embryology area. Faculty of Physiotherapy. Department of functional biology and health sciences, Pontevedra, Spain.

[15]Anhembi Morumbi University, School of Medicine, São José dos Campos, São Paulo, Brazil.

[16]Department of Dental Materials and Prosthodontics, São Paulo State University (Unesp), School of Dentistry, Araraquara, Brazil

[17]Oral Pathology, Department of Stomatology, Public Oral Health, and Forensic Dentistry, Ribeirão Preto Dental School (FORP/USP), University of São Paulo, Ribeirão Preto, SP 14040-904, Brazil.

***Correspondence:**

Mario Pérez-Sayáns PhD, DDS. Oral Medicine, Oral Surgery and Implantology Unit (MedOralRes). Faculty of Medicine and Dentistry Universidade de Santiago de Compostela. Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain. Address: Entrerríos s/n, Santiago de Compostela C.P. 15782 Spain. TLF: 0034626233504. E-mail: mario.perez@usc.es

ABSTRACT

Proliferative verrucous leukoplakia (PVL) is the oral potentially malignant disorder (OPMD) with the highest malignant transformation rate, reaching 72.4% over time. Due to his clinical behaviour, there is no treatment available at this time. Thus, prevention and early diagnosis are extremely necessary for a better prognosis. Therefore, this study aimed to analyse the PVL proteome and find diagnostic biomarkers with therapeutic potential. Using tissue and saliva of PVL, oral leukoplakia diagnoses and control patients a proteomics analysis was performed through LC-MS/MS label-free, qualitative and quantitative analyses were conducted to identify differentially expressed proteins. Potential biomarkers were prioritized and further investigated by immunohistochemistry (IHC) and Marking Intensity (MI) scan-analyses in tissue samples of PVL, OL and control from Brazil, Spain and Finland. The immune system, cell cycle, DNA regulation and apoptosis pathways, as well as the whole proteome, were significantly different in the PVL samples. Besides, CALR, GNB2L1, YWHA family were highly differentially expressed in PVL samples through the LC-MS/MS analyses. IHC and MI analysis showed that CALR

higher expression can differentiate the PVL from OL, while YWHAQ and GNB2L were higher in PVL samples, and also proved to be helpful to discriminate the PVL from controls. Altogether, the proposed biomarkers bring new tools to improve the early detection of PVL, and these biomarkers could provide the possibility of new treatments since there is evidence of the modulation of these proteins can enhance better the clinical outcomes.

1. INTRODUCTION

Proliferative verrucous leukoplakia (PVL) is defined as a "distinct and aggressive form of oral potentially malignant disorder" (OPMD)[1], understanding as a variation from oral leukoplakia (OL)[2–5]. Clinically appear as a recalcitrant non-homogenous leukoplakia affecting unique or multiple sites, which potentially involve more contiguous or non-contiguous areas over time[6–10]. PVL is the OPMD with the highest malignant transformation rate, recently was determined a cumulative rate of 49.5% (CI99%: 26.7%- 72.4%) between studies with 12 to 20 years of follow-up and a malignant transformation rate per year of 9.3%[11]. Moreover, due to the high potential of recurrence and reaching several sites of the mouth, before and once is transformed there is no treatment available, consequently, patients will have low survival rates[12–15]. Thus, prevention and early diagnosis are extremely necessary for better prognosis[2,5,11,16,17].

Despite the different criteria developed over time[2,18–21], PVL diagnosis is still controversial proper to the lack of pathognomonic clinical, microscopic and molecular features. Besides, OPMD diagnoses and risk assessment based only on clinical and histopathological criteria are poorly reproducible and have high inter-observer variability, while the usefulness of dysplasia subdivision has been questioned due to the large intra-observer and inter-observer variability regarding the presence/absence and grade of dysplasia assessment[5,22–24]. There is recent evidence reinforcing the lack of agreement for PVL diagnosis based on this method[25]. Altogether, there is

no standardized objective method yet available for early detection, prevention, and recognize the progression of the development of oral carcinoma from PVL lesions.

The use of molecular techniques could markedly improve the detection of changes that are not visible on histopathological analysis, improving the effectiveness of identifying patients with a higher risk to develop a carcinoma[7,16,26,27]. Today there is not a consensus about which biomarkers could be helpful to make PVL diagnoses and/or to treatment this pathology[28]. Furthermore, a deeper understanding of the molecular nature of PVL is essential to develop new diagnostic and therapeutic tools.

Saliva is a complex matrix, whose composition can vary depending on autonomic nervous system stimulation, circadian rhythm, habits, health-disease status, among others[29–31]. This has aroused great interest in using saliva as a method of diagnosis and control of different diseases, both systemic and oral cavity[30,31], due to the bountiful molecular content, lesser invasiveness, and ease extraction, as well as the low-cost involved[30]. Advances in the field of saliva have established its usefulness as a source of biomarkers comparable to blood, cerebrospinal fluid, pleural fluid, and urine[30–33], which would allow early screening of different diseases through "liquid biopsies"[31,32,34].

The increasing development and innovation have made it possible to improve the resolution of proteomics based on mass spectrometry, reaching more sensitive detections on tissue as well in saliva[32–37]. Different software makes possible the differential quantitative analysis of the proteome, allowing to explore clusters and nodes of different networks, as well as their biological, functional, and metabolic pathways[38–43]. Consequently, the search for markers in tissue and saliva is of paramount importance cause opens up new possibilities of diagnosis and possible therapeutic targets for preventing progression to oral carcinoma.

Therefore, this study aimed to analyse descriptively the PVL proteome from tissue and saliva in comparison with OL and find diagnostic biomarkers with therapeutic potential from epithelial

tissue as well as the possibility to apply these through the liquid biopsy. Also, investigate their correlation with clinical and histopathological features.

2. METHODS

2.1. Design and type of study

This study was conducted based on a descriptive case and control clinical model of retrospective analysis with prospective recruitment of patients. Also, was designed following the REMARK recommendations for tumour marker prognostic studies[44].

The study was developed between the Oral Medicine Clinical Center of Araraquara, Faculty of Dentistry, Sao Paulo State University (UNESP), and the Oral Medicine, Oral Surgery and Implantology Unit (MedOralRes), Faculty of Medicine and Dentistry, University of Santiago de Compostela, both had the approval of their local institutional ethics committees, under the numbers 34361814.9.0000.5416 and 2019/271, respectively.

Data was gathered from March 2014 to November 2019. Before undertaking the proceedings, the subjects' informed consent was recorded in writing following the Declaration of Helsinki and its subsequent amendments.

2.2. Target conditions and control

Patients included had been diagnosed with PVL and OL, according to the WHO Classification for Head and Neck Tumours[1], with different dysplasia grades without carcinoma *in situ* signals, for PVL patients were chosen specifically lesions with an epithelium with verrucous appearance. For control purposes were selected patients with an inflammatory reactional non-neoplasm process of the oral mucosa.

2.3. Other exclusion criteria

Were excluded from the study patients diagnosed with malignant conditions and others OPMD as lichen planus, oral lichenoid lesion, oral submucous fibrosis, erythroplakia, actinic cheilitis,

oral squamous cell carcinoma, erythroleukoplakia, nicotinic stomatitis, and also patients with lesions that showed histologically sign of carcinoma *in situ* or microinvasive carcinoma.

2.4. Subjects

Tissue and saliva samples were compound for 3 groups; PVL, OL, and inflammatory fibrous hyperplasia (IFH). For tissue, were five samples for each one group (1 sample from the IFH group was missing during protein extraction, so the group was reduced to 4 samples). Saliva groups were compound for 7 PVL, 8 OL, and 5 IFH samples. (Details of the patient demographic data and samples from each group are shown in Table 1 and supplementary Table 1).

2.5. Study variables

Both types of samples were included as variables: diagnosis, substrate, sex, date of birth, age, date of sample, date of initial diagnosis, race, clinical presentation, anatomic location of biopsy, number of lesions, type of biopsy, type of treatment performed, comorbidities, dysplasia grade, previous carcinoma, clinical evolution in the period studied, smoking, cigarettes per day, years of smoking, years of smoking ex-smoker and alcohol consumption.

2.6. Tissue samples processing for LC-MS/MS Qualitative (DDA) and Quantitative (Proteome Discoverer 2.1 by Sequest HT algorithm) analysis

2.6.1. Collection and preparation

Tissue fragments (20-25mm) were obtained by incisional biopsy, each one was divided into two parts of 10mm, one of these fragments was directed for routine histopathological analysis and another fragment to freeze for this study. The freezing process was performed with Tissue Tek O.C.T. compound in plastic cryomold and liquid nitrogen, followed by storage at -80ºC. The tissues for histopathological analysis were fixed in formalin 10% and processed with haematoxylin-eosin colouration, to confirm the clinical diagnoses.

2.6.2. Preparation of the slides and laser microdissection (LM)

The frozen samples were cut at 9μm thick slices in a portable cryostat (Leica Microsystem) at -23°C. Arcturus PEN Membrane Glass Slides (Life Technologies) was used to assemble the slices and were immediately stored in a plastic box at -80 °C[35]. Six slices of biopsy tissue were fixed in each slide, three slides with different profundity in the microanatomy were prepared for each sample in every group.

Laser microdissections of the slides containing the slices were coloured with toluidine blue staining to be microdissected in the Arcturus XT ™ IR-Laser (LCM - Laser Capture Microdissection - Life Technologies), each microdissected tissue was coupled to a cap, obtained 3 caps for every sample. The microdissected areas correspond just to the stratified squamous epithelial tissue.

2.6.3. Protein extraction and digestion

Caps with epithelial tissue were coupled in a tube to make a protein extraction and digestion[35]. First, urea 8M was added and then incubated with dithiothreitol (DTT) at 5mM for 25min at 56°C. Afterwards, iodoacetamide (IAA) was added at a final concentration of 14mM followed by incubation for 30 min in the dark at room temperature. The quench of free IAA was performed with 5mM DTT via incubation for 15min in the dark at room temperature. Finally, the samples were diluted 10-fold with 100mMtris pH8.0 and incubated with sequence grade modified Trypsin (Promega) at 1:50 concentration for trypsin: protein during 16hrs at 37°C.

Samples were acidified with formic acid. The resulting peptides were loaded onto a tip column Porous R2 analogue washed with trifluoroacetic acid (TFA) 0.4 %, with a pH low than 2.0. Fractions were collected and dried in a speed vacuum concentrator at 2500g for 10min. Dried peptides were dissolved in 0.1 % formic acid (FA)[35].

After the protein digestion, normalization was performed considering the lowest size area of epithelium adhered to the cap, which was determinate by an area of 9.046.307μm^2 that should

correspond to a final volume of 20µL. Following this rule were defined the necessary volume of FA to add in each sample of the three different groups, to obtain the same protein concentration relative to the lowest area of epithelium adhered to the cap before the LC-MS/MS analysis[35].

2.6.4. LC-MS/MS analysis

Each fraction was analysed in three technical replicates in an Easy-nLC 1000 nano-LC system (Thermo Scientific) coupled to a Q Exactive Plus Orbitrap mass spectrometer (Thermo Scientific). Samples were loaded onto a trap column (Thermo Scientific EASY-column C18, 2cm×100µm i.d.×5µm, 120Å) with a flow rate of 5µL/min and separated on the analytical column (Thermo Scientific EASY-column C18, 10cm×175µm i.d.×3µm) with a constant flow rate of 250nL/min and gradient of 5–45% for 100min, 45–95% for 7min of B (95% acetonitrile (ACN), 0.1% FA). For electrospray was used 2.7kV and 250°C at the inlet of the mass spectrometer. The instrument was operated in label-free data-dependent acquisition (DDA) mode with a dynamic exclusion of 45ms and full-scan MS spectra with a resolution of 70,000 followed by fragmentation of 15 most intense ions with HCD, NCE of 30, and resolution of 17,500 in MS/MS scans. Species with a charge of +1 and unassigned were excluded from MS/MS analysis[35].

2.6.5. Data analysis

Raw data were processed using Proteome Discoverer 2.1 Software (Thermo Scientific). Peptide identification was performed with the Sequest HT algorithm against the *Homo sapiens* database provided by Uniprot. The searches were performed with peptide mass tolerance of 10ppm, MS/MS tolerance of 0.05Da, tryptic cleavage specificity, 2 maximum missed cleavage sites, fixed modification of carbamidomethyl (Cys) and variable modification of acetylation of protein N-Terminus and oxidation of Methionine. False discovery rates (FDR) were obtained using

Percolator node selecting identifications with a q-value ≤ 0.01. (simplified workflow is shown on supplementary 1)

Statistical analysis was performed with Perseus 1.6.10.43 software, which is available in the MaxQuant package. Protein abundance, which was calculated based on ratio area/spectrum intensity, obtained from Proteome Discoverer was loaded on Perseus[38]. The workflow starts determining a categorical annotation for each sample and group, then the dataset was filtered by; filter rows based on valid values with minimum 70% valid values in at least one group of "*samples*". A general overview of exclusive and common proteins between the groups was performed by a Venn diagram using Bioinformatics & Evolutionary Genomics software.

Biological interaction pathways and biological processes were performed by Reactome[40] and FunRich[45,46] open-access software (Functional Enrichment analysis tool), respectively. For functional enrichment and interaction network analysis of the identified proteome (https://reactome.org, http://funrich.org/index.html).

Subsequently, the data were log2(x) transformed and applied z-score normalization of the samples by column using the median, then was calculated the median of the technical replicates followed by filter rows based on valid values with minimum 100% valid values in at least one group of "*group*"[35,39]. Over this last matrix, missing values for the LFQ intensity were imputed with random numbers from a normal distribution, the mean and standard deviation of which were selected to best simulate low abundance values close to the noise level (imputation width = 0.3, shift = 1.8)[35]. Then, a multivariate statistical analysis using principal component analysis (PCA) was performed to compare the data across the samples (cut-off method Benjamini-Hochberg FDR=0.05). Also, two volcano plots were performed to explore the foldchange (FC) values for protein expressions on the PLV group against OL and IFH group (using a t-test FDR=0.05 and S0=0.1)[35,39]. A threshold of ±1.2 for up/down expression was considered.

significant differential expression (DE) was assessed over the same imputed matrix using one-way ANOVA (Permutation-based FDR=0.05, with 250 randomizations) and Post Hoc test (FDR=0.05) to compare the three groups. Gene ontology (GO) of the biological process (GOBP), cellular component (GOCC), molecular function (GOMF) and pathways using Kyoto Encyclopaedia of Genes and Genomes (KEGG) data from *Homo sapiens* annotations of Uniprot database was added to perform enrichment theoretical analysis for these proteins.

For data visualization of clustering and expression of the proteins in the samples, heat maps with z-score values of log2 LFQ intensities were built selecting only proteins that showed DE values for OPMD lesions (potential biomarkers), as follow; PVL to OL and IFH, PVL to IFH and OL to IFH (simplified workflow is shown on supplementary Contextualtual network analysis by Cytoscape by CHAT was performed considering the proteins matrix of PVL/OL FC values (considering important the proteins with values >1.2) [41].

Proteins that showed significant DE values were loaded on STRING open-access software (using *homo sapiens* database, confidence network edges, with highest confidence value =0.900 and hiding disconnected nodes) to create a network to highlight the clusters and nodes formed by these proteins (https://string-db.org)[47,48]

The whole, the selection of the potential biomarkers for immunohistochemistry assays was made using the following criteria:

(I) Significant DE ($p<0.05$) between PVL against all groups. (II)Proteins were considered important from GO annotations when related to the *regulation of cell cycle, apoptosis, ageing, pathways in cancer, cellular replication, antigen processing and presentation, biological regulation, cell death*. (III)Theoretical analysis of the proteins highlighted as important in the previous points based on the literature research of processes related to malignant transformation and their therapeutic potential. (IV)The highest prevalence in the identification across the samples on their technical and biological replicates on tissue LC-MS/MS analysis. (V)"Up

expressed" proteins for difference values between PVL/OL, and also the behaviour in PVL/IFH and OL/IFH FC values. (VI)Consider the nodes that connect the biggest cluster and the possible effector nodes on the STRING network. (VII)Considering highlighted proteins and clusters on CHAT app.

2.7. Saliva samples processing for LC-MS/MS Qualitative (DDA) and Quantitative (SWATH[49]) analysis

2.7.1. Saliva Sample Collection and Processing

Unstimulated whole saliva (5mL) was collected from case and control subjects by having each subject swallow and then expectorate continuously into 20-ml sterile, polypropylene conical tube for a period of 5 to 10 minutes[50]. To minimize the effect of food intake and circadian variation, all samples were taken at 8–9 am after the nothing by mouth state, also in other to avoid blood contamination of the saliva, participants were asked not to brush their teeth within 45 min before sample collection; saliva samples visibly contaminated with blood were discarded and recollected[32].

Two aliquots of 30 µL of saliva are precipitated using the method adapted from the described by Dr Wessel's group[49,51]. The pellet is resuspended in milli-Qwater and submitted to an in-gel concentration and in-gel digestion.

Therefore, to make global protein identification, an equal amount of protein (90µg) from all samples were loaded on a 10% SDS-PAGE gel. The run was stopped as soon as the front had penetrated 3 mm into the resolving gel [52,53] The protein band was detected by Sypro-Ruby fluorescent staining (Lonza, Switzerland), excised, and processed for in-gel, manual tryptic digestion as described elsewhere[54]. Peptides were extracted by carrying out three 20-min incubations in 40µL of 60% acetonitrile dissolved in 0.5% HCOOH. The resulting peptide extracts were pooled, concentrated in a SpeedVac, and stored at −20 °C.

In order to make global protein identification, an equal amount of protein (90µg) from all samples were loaded on a 10% SDS-PAGE gel. The run was stopped as soon as the front had penetrated 3 mm into the resolving gel [52,53] The protein band was detected by Sypro-Ruby fluorescent staining (Lonza, Switzerland), excised, and processed for in-gel, manual tryptic digestion as described elsewhere[54]. Peptides were extracted by carrying out three 20-min incubations in 40µL of 60% acetonitrile dissolved in 0.5% HCOOH. The resulting peptide extracts were pooled, concentrated in a SpeedVac, and stored at −20 °C.

2.7.2. Mass spectrometric analysis

To perform the LC-MS/MS, samples were dissolved in mobile phase A (0.1% formic acid in water). 4µL (4µg) of digested peptides were separated using Reverse Phase Chromatography. The gradient was created using a micro liquid chromatography system (Eksigent Technologies nanoLC 400, SCIEX) coupled to a high-speed Triple TOF 6600 mass spectrometer (SCIEX) with a microflow source. The chosen analytical column was a silica-based reversed-phase column Chrom XP C18 150 × 0.30 mm, 3 mm particle size and 120 Å pore size (Eksigent, SCIEX). The trap column was a YMC-TRIART C18 (YMC Technologies, Teknokroma with a 3 mm particle size and 120 Å pore size, switched on-line with the analytical column. The loading pump delivered a solution of 0.1% formic acid in water at 10 µL/min. The micro-pump generated a flow-rate of 5 µl/min and was operated under gradient elution conditions, using 0.1% formic acid in water as mobile phase A, and 0.1% formic acid in acetonitrile as mobile phase B. Peptides were separated using a 90 minutes gradient ranging from 2% to 90% mobile phase B (mobile phase A: 2% acetonitrile, 0.1% formic acid; mobile phase B: 100% acetonitrile, 0.1% formic acid).

Data acquisition was performed in a TripleTOF 6600 System (SCIEX, Foster City, CA) using a Data-dependent workflow (DDA). Source and interface conditions were the following: ion spray

voltage floating (ISVF) 5500 V, curtain gas (CUR) 25, collision energy (CE) 10 and ion source gas 1 (GS1) 25. The instrument was operated with Analyst TF 1.7.1 software (SCIEX, USA). Switching criteria was set to ions greater than mass to charge ratio (m/z) 350 and smaller than m/z 1400 with a charge state of 2–5, mass tolerance 250ppm and an abundance threshold of more than 200 counts (cps). Former target ions were excluded for 15 s. The instrument was automatically calibrated every 4 hours using external calibrant tryptic peptides from PepCalMix.

2.7.3. Data Analysis

After MS/MS analysis, data files were processed using ProteinPilotTM 5.0.1 software from Sciex which uses the algorithm ParagonTM for database search and ProgroupTM for data grouping. Data were searched using a Human-specific Uniprot database. False discovery rate was performed using a non-linear fitting method displaying only those results that reported a1% Global false discovery rate (FDR) or better. [55]

2.7.4. Protein quantification by SWATH (Sequential Window Acquisition of all Theoretical Mass Spectra)

2.7.4.1. Creation of the spectral library

In order to construct the MS/MS spectral libraries, the peptide solutions were analysed by a shotgun data-dependent acquisition (DDA) approach by micro-LC-MS/MS, as described previously. To get a good representation of the peptides and proteins present in all samples, pooled vials of samples from each group (IFH, OL and PVL) were prepared using equal mixtures of the original samples. 4µL (4µg) of each pool was separated into a micro-LC system Ekspert nLC425 (Eksigen, Dublin, CA, USA) using a column Chrom XP C18 150 × 0.30 mm, 3 mm particle size and 120 Å pore size (Eksigent, SCIEX). at a flow rate of 5µL/min. Water and ACN, both containing 0.1% formic acid, were used as solvents A and B, respectively. The gradient run consisted of 5% to 95% B for 30 min, 5 min at 90% B and finally 5 min at 5% B for column

equilibration, for a total run time of 40 min. When the peptides were eluted, they were directly injected into a hybrid quadrupole-TOF mass spectrometer Triple TOF 6600 (Sciex, Redwood City, CA, USA) operated with a data-dependent acquisition system in positive ion mode. A Micro source (Sciex) was used for the interface between microLC and MS, with an application of 2600 V voltage. The acquisition mode consisted of a 250 ms survey MS scan from 400 to 1250 m/z followed by an MS/MS scan from 100 to 1500 m/z (25 ms acquisition time) of the top 65 precursor ions from the survey scan, for a total cycle time of 2.8 s. The fragmented precursors were then added to a dynamic exclusion list for 15 s; any singly charged ions were excluded from the MS/MS analysis.

The peptide and protein identifications were performed using Protein Pilot software (version 5.0.1, Sciex) with Data that were searched using a Human-specific Uniprot database, specifying iodoacetamide as Cys alkylation. The false discovery rate (FDR) was set to 1 for both peptides and proteins. The MS/MS spectra of the identified peptides were then used to generate the spectral library for SWATH peak extraction using the add-in for PeakView Software (version 2.2, Sciex) MS/MSALL with SWATH Acquisition MicroApp (version 2.0, Sciex). Peptides with a confidence score above 99% (as obtained from Protein Pilot database search were included in the spectral library).

2.7.4.2. Relative quantification by SWATH acquisition

SWATH (Sequential Window Acquisition of all Theoretical Mass Spectra) – MS acquisition was performed on a TripleTOF® 6600 LC-MS/MS system (AB SCIEX). Samples were analysed using a data-independent acquisition (DIA) method (30 total samples). Each sample (4 µL (from an mg/ml solution) was analysed using the LC-MS equipment and LC gradient described above for building the spectral library but instead using the SWATH-MS acquisition method. The method consisted of repeating a cycle that consisted of the acquisition of 100 TOF MS/MS scans

(400 to 1500 m/z, high sensitivity mode, 50 ms acquisition time) of overlapping sequential precursor isolation windows of variable width (1 m/z overlap) covering the 400 to 1250 m/z mass range with a previous TOF MS scan (400 to 1500 m/z, 50 ms acquisition time) for each cycle. The total cycle time was 6.3 s. For each sample set, the width of the 100 variable windows was optimized according to the ion density found in the DDA runs using a SWATH variable window calculator worksheet from Sciex.

2.7.4.3. Data analysis

The targeted data extraction of the fragment ion chromatogram traces from the SWATH runs was performed by PeakView (version 2.2) using the SWATH Acquisition MicroApp(version 2.0). This application processed the data using the spectral library created from the shotgun data. The retention times from the peptides that were selected for each protein were realigned in each run according to the iRT of peptides from each sample that were eluted along the whole-time axis. PeakView computed an FDR and a score for each assigned peptide according to the chromatographic and spectra components; only peptides with an FDR below 1% were used for protein quantitation. Up to ten peptides per protein and seven fragments per peptide were selected, based on signal intensity; any shared and modified peptides were excluded from the processing. Five-minute windows and 30 ppm widths were used to extract the ion chromatograms; SWATH quantitation was attempted for all proteins in the ion library that were identified by ProteinPilot with an FDR below 1%. The extracted ion chromatograms were then generated for each selected fragment ion; the peak areas for the peptides were obtained by summing the peak areas from the corresponding fragment ions.

The integrated peak areas (processed. mrkvw files from PeakView) were directly exported to the MarkerView software (AB SCIEX) for relative quantitative analysis. The export will generate three files containing quantitative information about individual ions, the summed intensity of

different ions for a particular peptide and the summed intensity of different peptides for a particular protein. MarkerView has been used for the analysis of SWATH-MS data reported in other proteomics studies[56–59] because of its data-independent method of quantitation. MarkerView uses processing algorithms that accurately find chromatographic and spectral peaks direct from the raw SWATH data. Data alignment by MarkerView compensates for minor variations in both mass and retention time values, ensuring that identical compounds in different samples are accurately compared to one another. To control for possible uneven sample loss across the different samples during the sample preparation process, we performed a global normalization based on the total sum of all the peak areas extracted from all the peptides and transitions across the replicates of each sample. Unsupervised multivariate statistical analysis using principal component analysis (PCA) was performed to compare the data across the samples (PCA were plotting using as parameters square root and range scale). The average MS peak area of each protein was derived from the replicates of the SWATH-MS of each sample followed by Student's *t*-test analysis using the MarkerView software for comparison among the samples based on the averaged area sums of all the transitions derived for each protein. The *t*-test will indicate how well each variable distinguishes the two groups, reported as a *p*-value.

A general overview of exclusive and common proteins between the groups was performed using the library of Protein Pilot database selectin proteins with FDR below 1% (by a Venn diagram using Bioinformatics & Evolutionary Genomics software).

Biological interaction pathways and biological processes were performed byReactome[40] and FunRich[45,46]open-access software (Functional Enrichment analysis tool), respectively (https://reactome.org,http://funrich.org/index.html).

For this library, was considered in each comparison that a protein was differentially expressed when has a p-value <0.05 and FC> 1.2-fold or FC< 0.83.

All proteins identification on PVL saliva samples and proteins that showed DE values were loaded separately on STRING open-access software(using *homo sapiens* database, confidence network edges, with highest confidence value =0.900 and hiding disconnected nodes) to create two different networks highlight the clusters and nodes formed by these proteins (https://string-db.org)[47,48]

2.8. Multiple comparisons

Considering the limitation in the number of samples used for the LC-MS/MS a combined analysis using tissue and saliva samples was performed to evaluate the intensity of protein identification in the most important proteins presented in tissue and saliva samples related to the grade of dysplasia.

2.9. Immunohistochemical methods

For immunohistochemical (IHC) staining, 3 μm thick sections were placed on slides properly coated with organo-silane (Sigma-Aldrich, St Louis, MO, USA), were included for this stage; 31 PVL samples from Spain, 31 PVL, 32 OL and 25 IFH cases from Brazil, also 75 PVL and 17 OL samples from Finland. Briefly, the sections were deparaffinized and rehydrated in graded ethanol solutions. After antigen retrieval with EDTA/Tris buffer (pH 9.0) in a microwave oven (1380 W; Panasonic, Sa˜o Paulo, Brazil), endogenous peroxidase activity was blocked with 20% H_2O_2 using five cycles of 5 min each. Overnight incubation with the primary antibodies for CALR, YWHAQ and GNB2L1 (Table 2) diluted in bovine serum albumin (BSA) was followed by incubation with the secondary antibody conjugated with polymer dextran marked with peroxidase (Dako EnVision Labeled Polymer; Dako, Glostrup, Denmark). Reactions were developed with a solution containing 0.6mg/ml 3,3'diaminobenzidine tetrahydrochloride (DAB, Sigma-Aldrich, Saint Louis, MO, USA) and 0.01% H_2O_2 and counterstained with Carazzi's haematoxylin for 5 minutes. Thereafter, the sections were dehydrated in a series of graded

ethanol solutions, diaphanized and mounted in Canada balsam under cover glasses. Positive and negative controls were included in all reactions.

2.10. Marking Intensity of Immunohistochemistry assay

Proteins expression was assessed with the aid of the Aperio ImageScope 12.4.3 (Leica Biosystems Division of Leica Microsystems Inc., USA). This step was performed by blind evaluators who didn't know which lesion was assessed. Briefly, glass slides were scanned into high-resolution images, which were analysed in the Pixel Count V9 algorithm software (Aperio ImageScope 12.4.3, USA). The keratinocyte cells were delimitated in a zoom x5 view and by using specific input parameters (hue value = 0.101, hue width = 0.00, colour saturation threshold = 0.101), the percentage of cytoplasm positivity was calculated and classified in three range categories, according to its staining intensity as weak (from 160 to 244), moderate (from 138 to 160) and strong (from 100 to 138). To each category, an intensity score was set: 1 for weak, 2 for moderate, and 3 for strong staining. Marking Intensity (MI) final scores were calculated as the sum of the percentage of each category weighted by its intensity score, using the following equation: [(%weak x 1) + (%moderate x 2) + (%strong x 3)][60,61]. Marking Intensity findings are described using descriptive statistics.

Statistical analysis was performed by a blinded statistician, without knowing who were the groups compared. Normal distribution was analysed using the Shapiro-Wilk test. The presence of outliers was checked. Homoscedasticity was assessed through Box's Test of Equality of Covariance Matrices and Levene's Test of Equality Variances. Therefore, the multivariate analysis of variance (Manova one-way) was applied. Differences among groups were analysed using the post hoc Games-Howell test. Statistical analysis was performed by IBM SPSS Statistics 20.0 (α= 0.05) and graphics images were built using GraphPad Prism version 6.0.

3. RESULTS

3.1. Clinical and demographic data

Table 1 summarises the characteristics of the patients and samples included in this study. The PVL population consist mainly of women (91.7%), while de OL was compound mostly by men (69.2%). PVL samples were biopsied principally from jugal mucosa, tongue and keratinized gingival tissue, as well, OL samples were biopsied mainly from the tongue (p=0.029). More than half of the PVL patients did not present associated comorbidities as well as the OL patients. PVL samples showed mainly high-grade (41.7%) and low-grade dysplasia (33.3%), while the OL samples exhibited low grade (53.8%) and absence (30.8%) of dysplasia. Over 83% of the population in OPMD the samples were non-smokers and no alcohol consumption was declared.

3.1.1. Follow-up in the study period

Till the redaction of this manuscript three of the five patients from the PVL tissue samples group developed an oncological event, and the other two presented recurrence in the PVL lesion without signs of malignancy in the histological analysis, there are still under follow up.

3.2. Protein Discoverer and overview of the proteome of tissue samples

The statistical analysis of detected proteins started with 663 proteins (FDR≤ 0.01) identified between groups from Proteome Discoverer (data are not shown). After the filtering process was kept 309 proteins, and the general overview of these by Venn diagram exhibited 110 and 2 exclusive proteins for PVL and OL, respectively (figure 1).

3.2.1. Biological Interaction Pathways and Biological process of tissue samples

Using the UniProt code of each protein of the PVL group, 266 out of 305 identifiers in the sample were found in Reactome, where 851 pathways were hit by at least one of them (details are shown in report 1, supplementary material), and the most significant represented pathways

were related to the immune system, cell cycle, DNA replication, apoptosis and metabolism of proteins (figure 2 & Supplementary 5). While in the OL group, 137 out of 157 identifiers in the sample were found in Reactome, where 652 pathways were hit by at least one of them (details are shown in report 2, supplementary material), and the most significant represented pathways were related to DNA repair, apoptosis, cellular responses to stress, immune system and cell cycle (figure 2 & Supplementary 6).

Biological processes, assessed by the FunRich tool; that involve the immune system, cell cycle control, programmed cell death, cell activity and the characteristics of cell phenotype have a higher number of proteins matched with those in the PVL group, in contrast, the to OL group (figure 3).

3.2.2. Proteome Analysis

Subsequently, the data related to the 309 proteins matrix were processed as mentioned above to develop a new matrix with 89 proteins.

3.2.2.1. Principal component analysis (PCA)

Using this last matrix, according to the principal components of the samples, the PVL group is located on the right side of the graph while the OL and IFH groups are located more nearly between them on the left side (PC1=36.9%, figure 4).

3.2.2.2. Significance Analysis of tissue samples

FC assessment of PLV/OL resulted in 32 significant proteins (figure 5) and PVL/IFH resulted in 37 significant proteins (figure 6).

DE was assessed and resulted in 51 differential expressed proteins between the groups (Permutation-based FDR=0.05, with 250 randomizations). GOBP, GOCC, GOMF and KEGG data from *Homo sapiens* annotations of the Uniprot database was added to perform enrichment theoretical analysis for these proteins.

Regarding OPMD samples only 43 proteins (potential biomarkers) showed significant DE values, post-test pointed up that 18 proteins of PVL in compression with OL and IFH at the same time, 10 proteins of PVL to OL, 13 proteins of PVL to IFH and 6 proteins of OL to IFH; were expressed significantly different (FDR=0.05) between the groups mentioned above (left-hand side of Table 3). Figure 7 shows the data visualization of clustering and expression of these potential biomarkers by a heatmap, this means that are displaying the intensities of each protein in every sample, allowing to build different clusters based on the amount of protein detected on the samples, note that the PVL samples are grouped strictly separate from the control group and the OL samples.

3.2.3. Bioinformatic functional network analysis

Contextual analysis performed over the 89 proteins matrix on Cytoscape by CHAT app shows that YWHA family proteins are weighty considering the FC values of PVL/OL (figure 8).

The STRING network built with the same matrix above for OPMD samples revealed six principal clusters, where one of the shows YWHA family proteins also linked with two other clusters compounded by chaperons' proteins like heat shock proteins, CALR and others like ALDOA, GAPDH, GDI2, RAB11A; another cluster is related to ribosomal protein activity, linked by GNB2L1, and the final is related to keratin proteins family (figure 9).

3.2.4. Selection of biomarkers

The whole, the selection of the potential biomarkers were made using the criteria mentioned in de material and methods section, the following proteins; identified on tissue samples, qualified as biomarkers and were chosen for immunohistochemistry assays: CALR, YWHAQ, GNB2L1.

3.3. Identification and overview of the proteome of saliva samples

Across the three groups, 374 proteins (FDR≤ 0.01) were identified from ProteinPilot™ (data are not shown). Between all these identified proteins; 282 corresponded to PVL, 287 to OL and 247

to the control group. To the general overview of the distribution of these identified proteins a Venn diagram was performed, exhibiting 190 common proteins in the groups and 53 and 49 exclusive proteins for PVL and OL, respectively (figure 10).

3.3.1. Biological Interaction Pathways and Biological process of saliva samples

Using the uniport code of each protein of the PVL group, 285 out of 369 identifiers in the sample were found in Reactome, where 1034 pathways were hit by at least one of them (details are shown in report 3, Supplementary 9), and the most significant represented pathways were related to the immune system and apoptosis (figure 11 & Supplementary 7). While in the OL group, 231 out of 287 identifiers in the sample were found in Reactome, where 908 pathways were hit by at least one of them (details are shown in report 4, Supplementary 10), and the most significant represented pathways were related to the immune system, DNA repair and cell cycle (figure 11 & Supplementary 8).

Biological processes that involve the immune system, cell cycle control, programmed cell death, cell activity and the characteristics of cell phenotype have a higher number of proteins linked with these in the PVL group, in contrast to the OL group (figure 12).

3.3.2. Proteome Analysis

3.3.2.1. Principal component analysis (PCA)

According to the principal components of saliva samples, the PVL group is located mainly on the left side while the OL and IFH groups are located more nearly between them on the right side in the graph (PC1=81.6%, figure 13).

3.3.2.2. Significance Analysis of tissue samples

DE and FC were assessed by MarkerView, this showed PLV/OL resulted in 60 significant proteins, with 28 proteins with FC values under the threshold (-1.2) and 12 proteins without statistical significance but they pointed up because their presence where detect on tissue

proteomic analysis (supplementary Table 2) and PVL/IFH resulted in 62 significant proteins, with 32 proteins with FC values under the threshold, 1 over the threshold and 13 proteins without statistical significance but they pointed up because their presence where detect on tissue proteomic analysis (supplementary Table 3).

3.4. Proteomic analysis of tissue and saliva samples

Based on DE and FC values, there were 20 proteins matched identified on tissue and saliva (top of Table 3). Nevertheless, each substrate had other proteins highlighted as important (bottom of Table 3); tissue and saliva samples presented 28 and 39 different proteins, respectively.

Table 4 presents multiple comparisons between protein expression and grade of dysplasia in all PVL samples analyzed by LC-MS/MS, note that almost of proteins showed a significantly higher amount of protein detected on high-grade dysplasia in comparison with low-grade and without dysplasia. Besides, the ENO1 and HSPA1B showed significant differences just in the comparison of high-grade and without dysplasia. Nonetheless, the comparison between the low-grade and without dysplasia only was significant for the SERPINB5 ($p=0.04$).

3.5. Marking Intensity of Immunohistochemistry assay

The independent factor assessed by each multivariate analysis was the biomarker (CALR, YWHAQ, GNB2L). The dependent variates were Group 1, Group 2, Group 3, which encodes the lesions groups OL, PVL, IFH, respectively. Thus, this characterized a one-way multivariate analysis of variance design. (The descriptive statistics are shown in Supplementary Table 4. An example of MI counting is shown in Supplementary 11).

The distribution of data was normal and heteroscedastic according to Box's Test of Equality of Covariance Matrices and Levene's Test of Equality Variances (both $p<0.001$). Therefore, Pillai's trace was used to evaluate the multivariate analysis of variance of the three biomarkers, showing statistically significant differences across the groups for each biomarker ($p<0.001$).

Games-Howell test showed statistically significant differences in the intraepithelial expression of CALR which was higher in PVL than OL ($p=0.003$), despite the expression of CALR being greater in PVL than IFH there was no statistically significant difference. The YWHAQ was broader expressed in OL compared with IFH ($p=0.002$) and also higher expressed in PVL in comparison with IFH ($p<0.001$), while the expression in PVL was higher than OL there was no statistically significant difference among these. Additionally, the GNB2L was more expressed in OL compared with IFH ($p<0.001$) and also greater in PVL in comparison with IFH ($p<0.001$), and was observed a higher expression in PVL than OL, however, there was no statistically significant difference between these (Statistical information is shown in Supplementary Tables 5 and 6).

4. DISCUSSION

Proteomics based on mass spectrometry is a promising approach for OPMD because allows the discovery of molecular signatures that clinical examination and histopathological analysis can't be seen [7,16,22–27]. Therefore, this study aimed to determine the PVL proteome from tissue and saliva in comparison with OL, and find diagnostic biomarkers with therapeutic potential from epithelial tissue as well as the possibility to apply these through the liquid biopsy, to develop new insight into diagnostic and treatment tools.

PVL samples showed a higher amount of protein detected in comparison with OL and controls in tissue samples as well in saliva (figure 1 and figure 10). Across the different pathways built with the identified proteins in PVL samples, both in tissue and saliva, is possible to note a huge difference between this lesion and OL regarding the immune system paths, certainly working in the immunosurveillance process[62–64], and this higher amount in immune system pathways would be explained probably due to the more malignant behaviour of PVL compared with OL[11] (figure 2 and figure 11).

Moreover, this greater malignancy capability would be related to a higher number of pathways from the cell cycle detected on PVL samples. Regardless, in response to the deregulation of the cell cycle, the cell can act also with another mechanism to control this situation, as was seen in the tissue PVL proteome; exhibiting a higher number of proteins related to apoptosis and DNA repair pathways (supplementary 5), or in saliva PVL proteome exhibiting a higher number of proteins related to apoptosis (supplementary 7). This molecular pattern seems to agree with the clinical behaviour of PVL, which shows a tendency to malignant transformation slowly and progressively over the time[1,2,10,12], without enough capacity of autoregulation since this lesion will progress to a carcinoma among 26.7% to 72.4% of cases[11].

Tissue and Saliva PCA (figure 4 and figure 13) shows clearly how is possible to separate and cluster the groups only based on the proteome of samples, PVL is further away from OL and control groups, demonstrating that these lesions have a different molecular composition, and the visualization of these features could help to solve the lack in the assessment based only in clinical and histopathological criteria's, which have proven poor reproducibility, high intra and inter-observer variability and conflicts in the agreement for PVL diagnosis[5,22–25]. These differences in the molecular composition in PVL suggest that this diagnosis could be a different entity from OL and should not be considered as a variation or more aggressive form of OL. Otherwise, the closeness between OL and control is in agreement with the potential of malignant transformation of OL, who not exceed 1.56% annually[11], since the control group are only inflammatory lesions without the potential of malignant transformation, it is logical to think that the samples that are closer to the control group (figure 4 and figure 13) will be more similar, molecularly, and consequently will have a lower potential for malignant transformation.

PVL not only present some exclusive proteins (figure 1 and figure 10), also tissue fold change assessment exposes that 32 proteins are present in PVL and OL and these are upper expressed in PVL samples (figure 6), as well 37 are upper expressed in PVL in comparison with the control

group (figure 7), demonstrating that the difference in clinical behaviour can be identified by the proteome of these (molecular pattern) and not only by histological analysis. The upper expression of these proteins also could explain the reason for PVL has the higher malignant transformation rate among OPMD[3–5,11,12] and the differences in the pathways expressed in PVL and OL (figure 2 , figure 11, supplementary 5, supplementary 6, supplementary 7 and supplementary 8).

Quantitative tissue analysis of DE reveals 43 proteins that are expressed significantly higher in PVL samples (FDR=0.05) (left-hand side of Table 3), and these lead to cluster the samples only based on intensity (figure 7), showing again that the proteome, in terms of differential protein expression, allows to differentiates the diagnoses.

The saliva proteome analysis showed a limited number of proteins identified also on tissue samples (top of Table 3), and not everyone showed significant DE or FC values, this lack could be due to the limited number used in LC-MS/MS analysis, but independently of this limitation, does not detract from the finding of the detection of important proteins in both, tissue and saliva, proven the utility of liquid biopsies.

Based on tissue PVL/OL FC values (left-hand side of Table 3) the contextual analysis highlighted the members of the YWHA proteins family, reflecting that the interactions on the different proteins that compound the matrix of PVL samples are strongly linked by the activity and function of these proteins (figure 8). Also, the network projected through STRING showed a cluster of the same proteins (figure 9), which have been implicated in the regulation of several intracellular signalling processes, including cell cycle, cell proliferation, cell migration, DNA damage checkpoint, apoptosis, autophagy, modulation of gene expression and regulation of oncoproteins and tumour suppressor proteins, in these last by binding YWHA proteins they can modulate their activity, cellular localization, stability and interactions[65–67]. Further, the

overexpression of YWHA proteins is strongly associated with development, poor survival rates and worst treatment outcomes in brain, lung, breast, liver and bladder tumours [66,68–80].

The YWHAQ also known as 14-3-3 protein theta/tau (14-3-3 θ/τ), is a member of the YWHA proteins family, who was proved *in vitro* assays when Tenascin-C, an extracellular matrix protein highly expressed in almost all solid tumours with antiadhesive properties, was expressed promoting growth of tumour cells by increasing 14-3-3τ expression, which consequently turns has a positive effect on tumour cell adhesion and growth[81]. Besides, studies in glioma showed that 14-3-3θ negatively regulates the nuclear function of the SLC2A4 regulator who induces cell apoptosis via caspase-3 and caspase-6 in glioma cells, leading to tumour progression[82]. Also, the 14-3-3 θ/τ protein suppressed the activity of signal-regulating kinase 1, which is responsible for inducing apoptosis in glioma cells[68].

In addition, other observations in cholangiocarcinoma (CCA) searching about notch signalling, who regulates a wide variety of processes like stem cell self-renewal, cell fate specification, differentiation, proliferation, and apoptosis; demonstrated by LC-MS/MS that NOTCH1 may cooperate with 14-3-3 theta to promote CCA cell survival[83]. Further, in lung cancer cells was detected an overexpression exclusively of 14-3-3 θ in comparison with normal mucosa, as well in breast cancer cells, in this last, the overexpression of 14-3-3 θ was also correlated with advanced TNM stage, lymph node metastasis, estrogenic negative status and poor clinical outcome [66,72,84]. Other studies showed that the reduction in 14-3-3 (109109ß, ε, ζ and θ) proteins whichs were overexpressed in cultured lung cells leads to a greater sensitivity to killing by radiation[71].

The Overexpression of 14-3-3τ in breast cancer cells exerts an inhibitory effect in tamoxifen-p21 induction and G1/S arrest leading to tumoral progression and then was also correlated with shorter patient survival[69]. Moreover, another study observed in breast cancer cell FC values over 1.5 for 14-3-3 θ/τ expression and this overexpression was significantly associated with

chemotherapy resistance[85]. Furthermore, a different study with western blot analysis of breast cancer cells demonstrated that the inhibition of 14-3-3τ with dicaffeoylquinic acids avoided cell proliferation and metastasis via Jak/PI3K/Akt and Raf/ERK pathway, which promote IL6 and CSF3 expression raised by CREB (CREBBP, CREB5) and induced cell apoptosis via Bad/Bax/caspase 9 signalling pathway[86].

In the LC-MS/MS performed in this study the YWHAQ was identified and differentially expressed on tissue ($q=0.02$) as well in saliva samples, but in this last almost reach significance ($q=0.06$) possibly due to the limited number of samples included in the LC-MS/MS. Otherwise, the FC values on tissue are opposite to the saliva samples, since saliva only contain secreted and non-linked forms of proteins, it is possible to think that the highest FC PVL/OL values on tissue represent the YWHAQ linked or acting in different biological process and pathways, while in saliva the PVL/OL FC value is too down expressed because much of the YWHAQ protein is interacting with other proteins and then the free/non-linked amount of YWHAQ is very limited in saliva. Despite that, the presence of this protein in PVL needs to be examined more deeply cause of the higher FC values of YWHA proteins in tissue samples, the opposite behaviour in FC values and also the detection of saliva samples is not a casualty (Table 3). Even more, taking all the m/z data of differentially expressed proteins the statistical analysis showed that the expression of YWHAQ, YWHAE, YWHAZ, YWHAB were significantly higher in high-grade dysplasia samples compared with low-grade and without dysplasia samples (Table 4). Moreover, the IHC assays analysed by MI showed higher YWHAQ expression on PVL samples than the other two groups, however only were statically difference between the OPMD lesions compared with control ($p<0.01$) and not between of them (supplementary Table 6) this information reinforce our idea that these proteins collaborate with the malignant transformation process of OPMD, mainly in PVL.

The YWHA cluster is linked with two clusters compounded mainly by chaperons' proteins like heat shock proteins (HSPA1B, HSPA8, HSP90AA1), CALR and others like ALDOA, GAPDH, GDI2, RAB11A (figure 9). Among these, heat shock proteins have the ability to maintain, stabilize and activated oncogenic proteins[87–89]. Indeed, HSP70 (encoded by HSPA1B) and HSP90 (encoded by HSP90AA1) balance p53 transcriptional behaviour, a well-known tumour suppressor, by regulating conformation and DNA binding activity[90]. Besides, HSP90 participates in the regulation of autophagy pathways which are implicated in cancer development, still, some inhibitors have been used for cancer therapy for preventing drug resistance, regulation of autophagy, as well as diminishing cellular proliferation, growth and inducing cell apoptosis[87–89,91–95]. Furthermore, HSP90 protein also was related to the promotion and survival of cells in Burkitt lymphoma[96] and nasopharyngeal carcinoma[95], and their depletion was correlated with better outcomes after surgery in gastric cancer[97]. On the other hand, some evidence released the possibility of bringing the immunotherapies with inhibitors of HSP90, since the inhibition can potentiate T-cell-mediated anti-tumour immune responses[98].

Calreticulin (CALR) is highly related to the carcinogenesis process. This protein acts as a cell surface marker for phagocytosis by macrophage or another immune cell, as the dendritic cell[99]. There is evidence showing that the macrophages would secret CALR to bind on asialoglycoproteins on the cell surface to target the cell and eliminate, and the malignant cell would express a significantly higher level of CALR-binding molecules than normal cells[100,101]. Besides, CALR would act favouriting the acetylation as an inhibitor histone deacetylase (HDAC) which would provide antitumor activity, upregulating genes responsible for apoptosis, cell cycle arrest and cellular differentiation and these hyperacetylation conditions also would improve enhance the therapeutic efficacy of the conventional antitumoral treatments[99,102,103].

Nevertheless, upper expression of CALR was associated with high tumour cell proliferation, cell growth, upregulation of vascular endothelial growth factor and metastasis, in some types of the

tumour as ovarian, pancreatic, gastric, prostate cancers and oral cancer, despite that, this relation does not mean causality between CALR and pro-tumoral cell behaviour[99,104]; regardless, contradictory observations are suggesting the deleterious effect of CALR in cancer cell survival and proliferation, like in bladder cancer, prostate cancers and neuroblastoma cells, where would diminish de epithelial-mesenchymal transition. Presently, is not clear if CALR exerts an effect prooncogenic, antitumoral or both depending on the location on cell and state (binding or solely) of this protein[99,105], or if his high presence in some type of cancers is a cellular response to stop the malignancy process, acting only as antitumoral protein.

The LC-MS/MS shows only the presence and differential expression of CALR on tissue samples ($q<0.00$) with high FC values over 1.2 to the PVL/OL and PVL/control comparisons (left-hand side of Table 3). The absence in saliva could be explained since CALR works as a protein-ligand and lives between the endoplasmic reticulum, cytosol and extracellular surface of the plasma membrane, when is secreted extracellularly will bind with other cell surfaces of plasma membrane[104,106–109]. On the other hand, the IHC assays analysed by MI showed higher CALR expression on PVL samples than the other groups, and only this biomarker was capable to distinguish between PVL and OL ($p<0.01$), although not between OPMD and control (supplementary Table 6), maybe due the inflammatory nature of the IFH.

Regarding the cluster with ribosomal proteins (RPs), linked by GNB2L1, could be hypostatized that they are in agreement with the functions of clusters mentioned above. Ribosomal proteins have extra-ribosomal functions, that can include participation in apoptosis, cell cycle arrest, cell proliferation, neoplastic transformation and cell migration and invasion[110]. The extra-ribosomal activity is chiefly mediated and regulated by the p53-MDM2 axis[110,111] or by the RPs binding through c-Myc and SP1, and nucleophosmin (p53-independent manner)[110,112–114].

The overexpression of RPs was associated with diverse types of cancers like prostate, gastric, lung, oesophageal, breast, osteosarcoma, renal cell carcinoma, melanoma, glioblastoma, ovarian

tumours; impacting on poor prognosis, short survival rates and negative predictor of metastasis-free survival[110]. Notwithstanding, in colorectal and liver cancer the overexpression seems to be associated with a better prognosis and better overall survival, respectively [110].

The Guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1) also known as Receptor for activated C kinase 1 (RACK1) in other species, is a multifunctional scaffold protein that acts in a wide range of biological processes, including signal transduction, immune response, cell growth, migration, cell adhesion and differentiation, and also has been implicated in the promotion of tumour invasion and metastasis[115–120]. Further, the inhibition of RACK1 affected negatively the cell proliferation, cell migration, invasion and adhesion in neuroblastoma[121], also induced apoptosis in hepatocellular carcinoma[122] and too provoked a reduction in senescence in cervical cancer cells[123].

The overexpression of RACK1 was reported in assorted kinds of tumours including breast, thyroid, melanoma, lung, brain, colorectal, pancreas and OSCC[124–134]. Regardless that the overexpression can promote progression in these tumours, in colon play an oncogenic role by the down expression, inducing autophagy and promoting proliferation and survival of tumour cells[135], furthermore, in lesions with helicobacter pylori infection, the downregulation leads to the elevated activation of the NF-κB signalling pathway, inducing carcinogenesis[136].

RACK1 stimulates the progression and survival of malignant tumours in OSCC through several pathways, including inhibition of G1 and G2 phase arresting via upregulation of Cyclin B1 and Cyclin D1. When RACK1 was knockdown led to decrease tumour volume and the expression of Ki67, CD34, and VEGF *in vivo*. Also, the knockdown can downregulate the protein levels of p-AKT, p-mTOR, and p-S6, affecting negatively growth, survival, adhesion, migration and invasion of tumoral cells[137,138]. The overexpression of RACK1 in OSCC was strongly related to Ki67, proving to be a good regulator of cell survival through favouring anti-apoptotic activities, also lesions with high expression of RACK1 were correlated with the severity of the epithelial

dysplasia, clinical stage, lymph node involvement, recurrence and more aggressive behaviour, the whole indicating a poor clinical outcome [133,139].

The LC-MS/MS shows only the presence and differential expression of GNB2L1on tissue samples ($q<0.00$) with high FC values over 1.2 to the PVL/OL and PVL/control comparisons (left-hand side of Table 3). The absence in saliva could be explained since GNB2L1 works as a protein receptor and is mainly located in the cell nucleus, perinuclear region, plasma membrane and rarely in the extracellular medium[140–144]. Furthermore, the IHC assays analysed by MI showed greater GNB2L1 expression on PVL samples than the other two groups, regardless of the statically difference only was between the PVL and OL cases compared with control ($p<0.01$) (supplementary Table 6). As well as in the case of the YWHAQ, it appears that these proteins only differentiate OPMD from controls but not from each other, maybe these two biomarkers are a common denominator in the malignant transformation process, and the amount expressed may be related to the magnitude of the malignant potential, as PVL samples always showed higher expression of these two biomarkers.

5. CONCLUSIONS

Despite that PVL is understood as a more aggressive form of OL, this work shows that these lesions have a different protein background, consequently, they should be considered as two different clinical entities, but which undoubtedly fall within the context of OPMD not only because of their clinical behaviour but also because of the protein machinery they possess, which facilitates the malignant transformation processes in different proportions depending on each lesion. Further, the nature of the behaviour of PVL is highly influenced by the YWHA family, CALR and GNB2L supporting the hypothesis that these proteins are related to the process of malignancy, either as its cause or as protectors of this event, depending on the specific function

of each protein. In addition, CALR proved to be useful as a biomarker to differentiate the PVL from OL, while YWHAQ and GNB2L proved to be helpful to discriminate the PVL and OL from controls.

Altogether, these findings need to be explored more deeply because they could provide new targets to develop diagnostics tools as well treatments to enhance the prognosis of patients with potentially malignant lesions.

ACKNOWLEDGEMENTS

To the Network of Multiuser Laboratories -RELAM/PROPP-UFU, aims to institutionalize equipment, facilities and technical staff, operating as a *facility.*

FUNDING: Esteban Arroyo received funding for this research from Estudantes-Convênio de Pós-Graduação – PEC-PG, CAPES/CNPq, Brazil (88881.154490/2017-01). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Andreia Bufalino received funding for this research from São Paulo Research Foundation (FAPESP) (grant #2017/01438-0) and the National Council for Scientific and Technological Development CNPq (423945/2016-5). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES).

COMPETING INTERESTS: The authors declare that they have no competing interests

REFERENCES

1. Oral potentially malignant disorders and oral epithelial dysplasia. in *WHO Classification of Head and Neck Tumours* (eds. El-Naggar, A. K., Chan, J. K., Grandis, J. R., Takata, T. & Slootweg, P. J.) 112–115 (International Agency for Research on Cancer, 2017).

2. Cerero-Lapiedra, R., Balade-Martinez, D., Moreno-Lopez, La., Esparza-Gomez, G. & Bagan, Jv. Proliferative verrucous leukoplakia: A proposal for diagnostic criteria. *Med. Oral Patol. Oral Cirugia Bucal* e839–e845 (2010) doi:10.4317/medoral.15.e839.

3. Van Der Waal, I. Potentially malignant disorders of the oral and oropharyngeal mucosa;

present concepts of management. *Oral Oncol.* **46**, 423–425 (2010).

4. Van der Waal, I. Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. *Oral Oncol.* **45**, 317–323 (2009).

5. Warnakulasuriya, S., Johnson, Newell. W. & Van Der Waal, I. Nomenclature and classification of potentially malignant disorders of the oral mucosa: Potentially malignant disorders. *J. Oral Pathol. Med.* **36**, 575–580 (2007).

6. Villa, A. *et al.* Proliferative leukoplakia: Proposed new clinical diagnostic criteria. *Oral Dis.* **24**, 749–760 (2018).

7. Celentano, A., Glurich, I., Borgnakke, W. S. & Farah, C. S. World Workshop on Oral Medicine VII: Prognostic biomarkers in oral leukoplakia and proliferative verrucous leukoplakia—A systematic review of retrospective studies. *Oral Dis.* odi.13363 (2020) doi:10.1111/odi.13363.

8. Capella, D. L., Gonçalves, J. M., Abrantes, A. A. A., Grando, L. J. & Daniel, F. I. Proliferative verrucous leukoplakia: diagnosis, management and current advances. *Braz. J. Otorhinolaryngol.* **83**, 585–593 (2017).

9. Borgna, S. C. *et al.* Management of proliferative verrucous leukoplakia: Justification for a conservative approach. *Head Neck* **39**, 1997–2003 (2017).

10. Torrejon-Moya, A., Jané-Salas, E. & López-López, J. Clinical manifestations of oral proliferative verrucous leukoplakia: A systematic review. *J. Oral Pathol. Med.* **49**, 404–408 (2020).

11. Iocca, O. *et al.* Potentially malignant disorders of the oral cavity and oral dysplasia: A systematic review and meta-analysis of malignant transformation rate by subtype. *Head Neck* **42**, 539–555 (2020).

12. Bagan, J. V. *et al.* Malignant transformation of proliferative verrucous leukoplakia to oral

squamous cell carcinoma: a series of 55 cases. *Oral Oncol.* **47**, 732–735 (2011).

13. Gouvea, Af., Moreira, Ae., Reis, Rr., Almeida, Op. & Lopes, Ma. Proliferative verrucous leukoplakia, squamous cell carcinoma and axillary metastasis. *Med. Oral Patol. Oral Cirugia Bucal* e704–e708 (2010) doi:10.4317/medoral.15.e704.

14. Scully, C. & Bagan, J. Oral squamous cell carcinoma: overview of current understanding of aetiopathogenesis and clinical implications. *Oral Dis.* **15**, 388–399 (2009).

15. Sklenicka, S., Gardiner, S., Dierks, E. J., Potter, B. E. & Bell, R. B. Survival Analysis and Risk Factors for Recurrence in Oral Squamous Cell Carcinoma: Does Surgical Salvage Affect Outcome? *J. Oral Maxillofac. Surg.* **68**, 1270–1275 (2010).

16. Farah, C. S. & Fox, S. A. Dysplastic oral leukoplakia is molecularly distinct from leukoplakia without dysplasia. *Oral Dis.* **25**, 1715–1723 (2019).

17. Villa, A. *et al.* Oral keratosis of unknown significance shares genomic overlap with oral dysplasia. *Oral Dis.* **25**, 1707–1714 (2019).

18. Hansen, L. S., Olson, J. A. & Silverman, S. Proliferative verrucous leukoplakia. A long-term study of thirty patients. *Oral Surg. Oral Med. Oral Pathol.* **60**, 285–298 (1985).

19. Ghazali, N., Bakri, M. & Zain, R. Aggressive, multifocal oral verrucous leukoplakia: Proliferative verrucous leukoplakia or not? *J. Oral Pathol. Med. Off. Publ. Int. Assoc. Oral Pathol. Am. Acad. Oral Pathol.* **32**, 383–92 (2003).

20. Bagan, J., Scully, C., Jimenez, Y. & Martorell, M. Proliferative verrucous leukoplakia: a concise update. *Oral Dis.* **16**, 328–332 (2010).

21. Carrard, V. C., Brouns, E. R. E. A. & van der Waal, I. Proliferative verrucous leukoplakia; a critical appraisal of the diagnostic criteria. *Med. Oral Patol. Oral Cir. Bucal* **18**, e411–e413 (2013).

22. Kujan, O. *et al.* Evaluation of a new binary system of grading oral epithelial dysplasia for prediction of malignant transformation. *Oral Oncol.* **42**, 987–993 (2006).

23. Fleskens, S. A. J. H. M. *et al.* Interobserver variability of laryngeal mucosal premalignant lesions: a histopathological evaluation. *Mod. Pathol.* **24**, 892–898 (2011).

24. Sperandio, M. *et al.* Predictive Value of Dysplasia Grading and DNA Ploidy in Malignant Transformation of Oral Potentially Malignant Disorders. *Cancer Prev. Res. (Phila. Pa.)* **6**, 822–831 (2013).

25. Upadhyaya, J. D. *et al.* Inter-observer Variability in the Diagnosis of Proliferative Verrucous Leukoplakia: Clinical Implications for Oral and Maxillofacial Surgeon Understanding: A Collaborative Pilot Study. *Head Neck Pathol.* **14**, 156–165 (2020).

26. Joseph, B. K. Oral Cancer: Prevention and Detection. *Oral Cancer* 4.

27. Farah, C. S. *et al.* Exome sequencing of oral leukoplakia and oral squamous cell carcinoma implicates DNA damage repair gene defects in malignant transformation. *Oral Oncol.* **96**, 42–50 (2019).

28. Rintala, M., Vahlberg, T., Salo, T. & Rautava, J. Proliferative verrucous leukoplakia and its tumor markers: Systematic review and meta-analysis. *Head Neck* **41**, 1499–1507 (2019).

29. Kaufman, E. & Lamster, I. B. T HE D IAGNOSTIC A PPLICATIONS OF S ALIVA — A R EVIEW. *Crit. Rev. Oral Biol. Med.* **13**, 197–212 (2002).

30. Gröschl, M. Saliva: a reliable sample matrix in bioanalytics. *Bioanalysis* **9**, 655–668 (2017).

31. Aro, K., Wei, F., Wong, D. T. & Tu, M. Saliva Liquid Biopsy for Point-of-Care Applications. *Front. Public Health* **5**, (2017).

32. Lorenzo-Pouso, A. I. *et al.* Protein-Based Salivary Profiles as Novel Biomarkers for Oral Diseases. *Dis. Markers* **2018**, 1–22 (2018).

33. Amado, F., Lobo, M. J. C., Domingues, P., Duarte, J. A. & Vitorino, R. Salivary peptidomics. *Expert Rev. Proteomics* **7**, 709–721 (2010).

34. Siravegna, G., Marsoni, S., Siena, S. & Bardelli, A. Integrating liquid biopsies into the

management of cancer. *Nat. Rev. Clin. Oncol.* **14**, 531–548 (2017).

35. Carnielli, C. M. *et al.* Combining discovery and targeted proteomics reveals a prognostic signature in oral cancer. *Nat. Commun.* **9**, 3598 (2018).

36. Johann, D. J. *et al.* Approaching Solid Tumor Heterogeneity on a Cellular Basis by Tissue Proteomics Using Laser Capture Microdissection and Biological Mass Spectrometry †. *J. Proteome Res.* **8**, 2310–2318 (2009).

37. Indovina, P. *et al.* Mass spectrometry-based proteomics: The road to lung cancer biomarker discovery. *Mass Spectrom. Rev.* **32**, 129–142 (2013).

38. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).

39. Tyanova, S. & Cox, J. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. in *Cancer Systems Biology* (ed. von Stechow, L.) vol. 1711 133–148 (Springer New York, 2018).

40. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).

41. Muetze, T. *et al.* Contextual Hub Analysis Tool (CHAT): A Cytoscape app for identifying contextually relevant hubs in biological networks. (2016).

42. Kuhn, M. *et al.* STITCH 4: integration of protein–chemical interactions with user data. *Nucleic Acids Res.* **42**, D401–D407 (2014).

43. Szklarczyk, D. *et al.* STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Res.* **44**, D380–D384 (2016).

44. Altman, D. G., McShane, L. M., Sauerbrei, W. & Taube, S. E. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *BMC Med.* **10**, 51 (2012).

45. Pathan, M. *et al.* FunRich: An open access standalone functional enrichment and

interaction network analysis tool. *PROTEOMICS* **15**, 2597–2601 (2015).

46. Pathan, M. *et al.* A novel community driven software for functional enrichment analysis of extracellular vesicles data. *J. Extracell. Vesicles* **6**, 1321455 (2017).

47. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).

48. Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).

49. Mdp, C.-V., M, G. V., A, S., C, N. & Sb, B. SWATH-MS Protocols in Human Diseases. *Methods Mol. Biol. Clifton NJ* **2259**, 105–141 (2021).

50. Navazesh, M. Methods for Collecting Saliva. *Ann. N. Y. Acad. Sci.* **694**, 72–77 (1993).

51. Wessel, D. & Flügge, U. I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **138**, 141–143 (1984).

52. Bonzon-Kulichenko, E. *et al.* A robust method for quantitative high-throughput analysis of proteomes by 18O labeling. *Mol. Cell. Proteomics MCP* **10**, M110.003335 (2011).

53. Perez-Hernandez, D. *et al.* The intracellular interactome of tetraspanin-enriched microdomains reveals their function as sorting machineries toward exosomes. *J. Biol. Chem.* **288**, 11649–61 (2013).

54. Shevchenko, A. *et al.* A strategy for identifying gel-separated proteins in sequence databases by MS alone. *Biochem Soc Trans* **24**, 893–896 (1996).

55. Shilov, I. V. *et al.* The Paragon Algorithm, a Next Generation Search Engine That Uses Sequence Temperature Values and Feature Probabilities to Identify Peptides from Tandem Mass Spectra. *Mol. Cell. Proteomics* **6**, 1638–1655 (2007).

56. Ortea, I., Ruiz-Sánchez, I., Cañete, R., Caballero-Villarraso, J. & Cañete, M. D. Identification of candidate serum biomarkers of childhood-onset growth hormone deficiency

using SWATH-MS and feature selection. *J. Proteomics* **175**, 105–113 (2018).

57. Meyer, J. G. & Schilling, B. Clinical applications of quantitative proteomics using targeted and untargeted data-independent acquisition techniques. *Expert Rev. Proteomics* **14**, 419–429 (2017).

58. Luo, Y. *et al.* SWATH-based proteomics identified carbonic anhydrase 2 as a potential diagnosis biomarker for nasopharyngeal carcinoma. *Sci. Rep.* **7**, 1–11 (2017).

59. Tan, H. T. & Chung, M. C. M. Label-Free Quantitative Phosphoproteomics Reveals Regulation of Vasodilator-Stimulated Phosphoprotein upon Stathmin-1 Silencing in a Pair of Isogenic Colorectal Cancer Cell Lines. *Proteomics* **18**, 1–30 (2018).

60. Bufalino, A. *et al.* Low miR-143/miR-145 Cluster Levels Induce Activin A Overexpression in Oral Squamous Cell Carcinomas, Which Contributes to Poor Prognosis. *PLOS ONE* **10**, e0136599 (2015).

61. Chang, K.-P. *et al.* Overexpression of activin A in oral squamous cell carcinoma: association with poor prognosis and tumor progression. *Ann. Surg. Oncol.* **17**, 1945–1956 (2010).

62. Zitvogel, L., Tesniere, A. & Kroemer, G. Cancer despite immunosurveillance: immunoselection and immunosubversion. *Nat. Rev. Immunol.* **6**, 715–727 (2006).

63. Dunn, G. P., Old, L. J. & Schreiber, R. D. The immunobiology of cancer immunosurveillance and immunoediting. *Immunity* **21**, 137–148 (2004).

64. Dunn, G. P., Old, L. J. & Schreiber, R. D. The Three Es of Cancer Immunoediting. *Annu. Rev. Immunol.* **22**, 329–360 (2004).

65. Gardino, A. K. & Yaffe, M. B. 14-3-3 proteins as signaling integration points for cell cycle control and apoptosis. *Semin. Cell Dev. Biol.* **22**, 688–695 (2011).

66. Khorrami, A., Sharif Bagheri, M., Tavallaei, M. & Gharechahi, J. The functional significance of 14-3-3 proteins in cancer: focus on lung cancer. *Horm. Mol. Biol. Clin. Investig.*

32, (2017).

67. Tzivion, G., Gupta, V. S., Kaplun, L. & Balan, V. 14-3-3 proteins as potential oncogenes. *Semin. Cancer Biol.* **16**, 203–213 (2006).

68. Yan, Y. *et al.* Implication of 14-3-3ε and 14-3-3θ/τ in proteasome inhibition-induced apoptosis of glioma cells. *Cancer Sci.* **104**, 55–61 (2013).

69. Wang, B. *et al.* 14-3-3τ Regulates Ubiquitin-Independent Proteasomal Degradation of p21, a Novel Mechanism of p21 Downregulation in Breast Cancer. *Mol. Cell. Biol.* **30**, 1508–1527 (2010).

70. Li, N. *et al.* Overexpression of 14-3-3θ promotes tumor metastasis and indicates poor prognosis in breast carcinoma. *Oncotarget* **5**, 249–257 (2014).

71. Qi, W. & Martinez, J. D. Reduction of 14-3-3 Proteins Correlates with Increased Sensitivity to Killing of Human Lung Cancer Cells by Ionizing Radiation. *Radiat. Res.* **160**, 217–223 (2003).

72. Qi, W., Liu, X., Qiao, D. & Martinez, J. D. Isoform-specific expression of 14-3-3 proteins in human lung cancer tissues. *Int. J. Cancer* **113**, 359–363 (2005).

73. Heidenblad, M. *et al.* Tiling resolution array CGH and high density expression profiling of urothelial carcinomas delineate genomic amplicons and candidate target genes specific for advanced tumors. *BMC Med. Genomics* **1**, 3 (2008).

74. Wu, C.-Y. *et al.* Prognostic Significance of 14-3-3ε, Aldo-keto Reductase Family 1 B10 and Metallothionein-1 in Hepatocellular Carcinoma. *Anticancer Res.* **38**, 6855–6863 (2018).

75. Tang, S. *et al.* 14-3-3ε Mediates the Cell Fate Decision-Making Pathways in Response of Hepatocellular Carcinoma to Bleomycin-Induced DNA Damage. *PLoS ONE* **8**, e55268 (2013).

76. Liu, T.-A. *et al.* 14-3-3ε Overexpression Contributes to Epithelial-Mesenchymal Transition of Hepatocellular Carcinoma. *PLoS ONE* **8**, e57968 (2013).

77. Ko, B.-S. *et al.* Overexpression of 14-3-3ε predicts tumour metastasis and poor survival

in hepatocellular carcinoma: Overexpression of 14-3-3ε in HCC. *Histopathology* **58**, 705–711 (2011).

78. Feng, Y.-H., Tung, C.-L., Su, Y.-C., Tsao, C.-J. & Wu, T.-F. Proteomic Profile of Sorafenib Resistance in Hepatocellular Carcinoma; GRP78 Expression Is Associated With Inferior Response to Sorafenib. *Cancer Genomics - Proteomics* **16**, 569–576 (2019).

79. Reis, H. *et al.* A structured proteomic approach identifies 14-3-3Sigma as a novel and reliable protein biomarker in panel based differential diagnostics of liver tumors. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1854**, 641–650 (2015).

80. Zhao, Y. *et al.* 14-3-3ζ/TGFβR1 promotes tumor metastasis in lung squamous cell carcinoma. *Oncotarget* **7**, 82972–82984 (2016).

81. Martin, D., Brown-Luedi, M. & Chiquet-Ehrismann, R. Tenascin-C signaling through induction of 14-3-3 tau. *J. Cell Biol.* **160**, 171–175 (2003).

82. Yun, D. *et al.* Shuttling SLC2A4RG is regulated by 14-3-3θ to modulate cell survival via caspase-3 and caspase-6 in human glioma. *EBioMedicine* **40**, 163–175 (2019).

83. Singrang, N., Kittisenachai, S., Roytrakul, S., Svasti, J. & Kangsamaksin, T. NOTCH1 regulates the viability of cholangiocarcinoma cells via 14-3-3 theta. *J. Cell Commun. Signal.* **13**, 245–254 (2019).

84. Li, N. *et al.* Overexpression of 14-3-3θ promotes tumor metastasis and indicates poor prognosis in breast carcinoma. *Oncotarget* **5**, 249–257 (2014).

85. Hodgkinson, V. C. *et al.* Proteomic identification of predictive biomarkers of resistance to neoadjuvant chemotherapy in luminal breast cancer: a possible role for 14-3-3 theta/tau and tBID? *J. Proteomics* **75**, 1276–1283 (2012).

86. Zhou, Y. *et al.* 1,3-Dicaffeoylquinic acid targeting 14-3-3 tau suppresses human breast cancer cell proliferation and metastasis through IL6/JAK2/PI3K pathway. *Biochem. Pharmacol.* **172**, 113752 (2020).

87. Workman, P., Burrows, F., Neckers, L. & Rosen, N. Drugging the Cancer Chaperone HSP90. *Ann. N. Y. Acad. Sci.* **1113**, 202–216 (2007).

88. Zuehlke, A. D., Beebe, K., Neckers, L. & Prince, T. Regulation and function of the human HSP90AA1 gene. *Gene* **570**, 8–16 (2015).

89. Calderwood, S. K. & Gong, J. Heat Shock Proteins Promote Cancer: It's a Protection Racket. *Trends Biochem. Sci.* **41**, 311–323 (2016).

90. Boysen, M., Kityk, R. & Mayer, M. P. Hsp70- and Hsp90-Mediated Regulation of the Conformation of p53 DNA Binding Domain and p53 Cancer Variants. *Mol. Cell* **74**, 831-843.e4 (2019).

91. Wang, B. *et al.* Hsp90 regulates autophagy and plays a role in cancer therapy. *Tumor Biol.* **37**, 1–6 (2016).

92. Cheng, W. *et al.* Role of the novel HSP90 inhibitor AUY922 in hepatocellular carcinoma: Potential for therapy. *Mol. Med. Rep.* **12**, 2451–2456 (2015).

93. Han, J. *et al.* HSP90 inhibition targets autophagy and induces a CASP9-dependent resistance mechanism in NSCLC. *Autophagy* 1–14 (2018) doi:10.1080/15548627.2018.1434471.

94. Shirota, T. *et al.* Heat Shock Protein 90 Is a Potential Therapeutic Target in Cholangiocarcinoma. *Mol. Cancer Ther.* **14**, 1985–1993 (2015).

95. Liu, F. *et al.* Clinical and biological significances of heat shock protein 90 (Hsp90) in human nasopharyngeal carcinoma cells and anti-cancer effects of Hsp90 inhibitor. *Biomed. Pharmacother.* **120**, 109533 (2019).

96. Walter, R. *et al.* HSP90 promotes Burkitt lymphoma cell survival by maintaining tonic B-cell receptor signaling. *Blood* **129**, 598–608 (2017).

97. Buffart, T. E. *et al.* Losses of chromosome 5q and 14q are associated with favorable clinical outcome of patients with gastric cancer. *The Oncologist* **17**, 653–662 (2012).

98. Mbofung, R. M. *et al.* HSP90 inhibition enhances cancer immunotherapy by upregulating

interferon response genes. *Nat. Commun.* **8**, 451 (2017).

99. Venkateswaran, K. *et al.* Emerging Roles of Calreticulin in Cancer: Implications for Therapy. *Curr. Protein Pept. Sci.* **19**, 344–357 (2018).

100. Feng, M. *et al.* Programmed cell removal by calreticulin in tissue homeostasis and cancer. *Nat. Commun.* **9**, 3194 (2018).

101. Duo, C.-C. *et al.* Soluble Calreticulin Induces Tumor Necrosis Factor-α (TNF-α) and Interleukin (IL)-6 Production by Macrophages through Mitogen-Activated Protein Kinase (MAPK) and NFκB Signaling Pathways. *Int. J. Mol. Sci.* **15**, 2916–2928 (2014).

102. Dwarakanath, B. S., Verma, A., Bhatt, A. N., Parmar, V. S. & Raj, H. G. Targeting protein acetylation for improving cancer therapy. *INDIAN J MED RES* 9 (2008).

103. Ponnan, P. *et al.* Comparison of Protein Acetyltransferase Action of CRTAase with the Prototypes of HAT. *Sci. World J.* **2014**, 1–9 (2014).

104. Chiang, W.-F. *et al.* Calreticulin, an endoplasmic reticulum-resident protein, is highly expressed and essential for cell proliferation and migration in oral squamous cell carcinoma. *Oral Oncol.* **49**, 534–541 (2013).

105. Houen, G. Calreticulin – Oncogene, Anti-oncogene, or Both? 2 (2019).

106. Arosa, F. A., de Jesus, O., Porto, G., Carmo, A. M. & de Sousa, M. Calreticulin Is Expressed on the Cell Surface of Activated Human Peripheral Blood T Lymphocytes in Association with Major Histocompatibility Complex Class I Molecules. *J. Biol. Chem.* **274**, 16917–16922 (1999).

107. Dupuis, M., Schaerer, E., Krause, K. H. & Tschopp, J. The calcium-binding protein calreticulin is a major constituent of lytic granules in cytolytic T lymphocytes. *J. Exp. Med.* **177**, 1–7 (1993).

108. Holaska, J. M. *et al.* Calreticulin Is a Receptor for Nuclear Export. *J. Cell Biol.* **152**, 127–140 (2001).

109. Gelebart, P., Opas, M. & Michalak, M. Calreticulin, a Ca2+-binding chaperone of the endoplasmic reticulum. *Int. J. Biochem. Cell Biol.* **37**, 260–266 (2005).

110. Xu, X., Xiong, X. & Sun, Y. The role of ribosomal proteins in the regulation of cell proliferation, tumorigenesis, and genomic integrity. *Sci. China Life Sci.* **59**, 656–672 (2016).

111. Zhou, X., Liao, J.-M., Liao, W.-J. & Lu, H. Scission of the p53-MDM2 Loop by Ribosomal Proteins. *Genes Cancer* **3**, 298–310 (2012).

112. Donati, G., Montanaro, L. & Derenzini, M. Ribosome biogenesis and control of cell proliferation: p53 is not alone. *Cancer Res.* **72**, 1602–1607 (2012).

113. Russo, A. *et al.* Human rpL3 induces G_1/S arrest or apoptosis by modulating p21 (waf1/cip1) levels in a p53-independent manner. *Cell Cycle Georget. Tex* **12**, 76–87 (2013).

114. Wanzel, M. *et al.* A ribosomal protein L23-nucleophosmin circuit coordinates Mizl function with cell growth. *Nat. Cell Biol.* **10**, 1051–1061 (2008).

115. Del Vecchio, I. *et al.* "Functional mapping of the promoter region of the GNB2L1 human gene coding for RACK1 scaffold protein". *Gene* **430**, 17–29 (2009).

116. Li, J.-J. & Xie, D. RACK1, a versatile hub in cancer. *Oncogene* **34**, 1890–1898 (2015).

117. Doan, A. T. & Huttenlocher, A. RACK1 regulates Src activity and modulates paxillin dynamics during cell migration. *Exp. Cell Res.* **313**, 2667–2679 (2007).

118. Gandin, V., Senft, D., Topisirovic, I. & Ronai, Z. A. RACK1 Function in Cell Motility and Protein Synthesis. *Genes Cancer* **4**, 369–377 (2013).

119. Ron, D. *et al.* RACK1 to the future--a historical perspective. *Cell Commun. Signal. CCS* **11**, 53 (2013).

120. Duff, D. & Long, A. Roles for RACK1 in cancer cell migration and invasion. *Cell. Signal.* **35**, 250–255 (2017).

121. Lu, F., Zhang, C., Wu, W.-J. & Wu, Y.-M. RACK1 downregulation suppresses migration and proliferation of neuroblastoma cell lines. *Oncol. Rep.* **27**, 1646–1652 (2012).

122. Zou, Y., Li, X., Zhang, Q. & Liu, D. RACK1 Silencing Induces Cell Apoptosis and Inhibits Cell Proliferation in Hepatocellular Carcinoma MHCC97-H Cells. *Pathol. Oncol. Res.* **24**, 101–107 (2018).

123. Liao, S. *et al.* The receptor for activated protein kinase C promotes cell growth, invasion and migration in cervical cancer. *Int. J. Oncol.* **51**, 1497–1507 (2017).

124. Cao, X.-X. *et al.* RACK1 promotes breast carcinoma migration/metastasis via activation of the RhoA/Rho kinase pathway. *Breast Cancer Res. Treat.* **126**, 555–563 (2011).

125. Cao, X.-X. *et al.* RACK1: A superior independent predictor for poor clinical outcome in breast cancer. *Int. J. Cancer* **127**, 1172–1179 (2010).

126. Myklebust, L. M., Akslen, L. A., Varhaug, J. E. & Lillehaug, J. R. Receptor for activated protein C kinase 1 (RACK1) is overexpressed in papillary thyroid carcinoma. *Thyroid Off. J. Am. Thyroid Assoc.* **21**, 1217–1225 (2011).

127. Egidy, G. *et al.* Transcription analysis in the MeLiM swine model identifies RACK1 as a potential marker of malignancy for human melanocytic proliferation. *Mol. Cancer* **7**, 34 (2008).

128. Zhong, X. *et al.* Overexpressions of RACK1 and CD147 associated with poor prognosis in stage T1 pulmonary adenocarcinoma. *Ann. Surg. Oncol.* **20**, 1044–1052 (2013).

129. Jin, S. *et al.* Overexpressed RACK1 is positively correlated with malignant degree of human colorectal carcinoma. *Mol. Biol. Rep.* **41**, 3393–3399 (2014).

130. Nagashio, R. *et al.* Expression of RACK1 is a novel biomarker in pulmonary adenocarcinomas. *Lung Cancer Amst. Neth.* **69**, 54–59 (2010).

131. Li, X. *et al.* RACK1 overexpression associates with pancreatic ductal adenocarcinoma growth and poor prognosis. *Exp. Mol. Pathol.* **101**, 176–186 (2016).

132. Lv, Q.-L. *et al.* Overexpression of RACK1 Promotes Metastasis by Enhancing Epithelial-Mesenchymal Transition and Predicts Poor Prognosis in Human Glioma. *Int. J. Environ. Res. Public. Health* **13**, (2016).

133. Wang, Z. *et al.* RACK1, an excellent predictor for poor clinical outcome in oral squamous carcinoma, similar to Ki67. *Eur. J. Cancer Oxf. Engl. 1990* **45**, 490–496 (2009).

134. Liu, S. *et al.* RACK1 is an organ-specific prognostic predictor in OSCC. *Oral Oncol.* **76**, 22–26 (2018).

135. Xiao, T. *et al.* RACK1 promotes tumorigenicity of colon cancer by inducing cell autophagy. *Cell Death Dis.* **9**, 1148 (2018).

136. Hu, Y. *et al.* Downregulation of tumor suppressor RACK1 by Helicobacter pylori infection promotes gastric carcinogenesis through the integrin β-1/NF-κB signaling pathway. *Cancer Lett.* **450**, 144–154 (2019).

137. Zhang, X. *et al.* Receptor for activated C kinase 1 (RACK1) promotes the progression of OSCC via the AKT/mTOR pathway. *Int. J. Oncol.* **49**, 539–548 (2016).

138. Li, J. *et al.* Receptor for activated C kinase 1 (RACK1): a regulator for migration and invasion in oral squamous cell carcinoma cells. *J. Cancer Res. Clin. Oncol.* **138**, 563–571 (2012).

139. Wang, Z. *et al.* Comparative Proteomics Approach to Screening of Potential Diagnostic and Therapeutic Targets for Oral Squamous Cell Carcinoma. *Mol. Cell. Proteomics MCP* **7**, 1639–1650 (2008).

140. Baumann, M. *et al.* The PKC targeting protein RACK1 interacts with the Epstein-Barr virus activator protein BZLF1: RACK1 and BZLF1 interact. *Eur. J. Biochem.* **267**, 3891–3901 (2000).

141. Gallina, A., Rossi, F. & Milanesi, G. Rack1 Binds HIV-1 Nef and Can Act as a Nef–Protein Kinase C Adaptor. *Virology* **283**, 7–18 (2001).

142. Chuang, N.-N. & Huang, C.-C. Interaction of integrin β1 with cytokeratin 1 in neuroblastoma NMB7 cells. *Biochem. Soc. Trans.* **35**, 1292–1294 (2007).

143. Chang, B. Y., Chiang, M. & Cartwright, C. A. The Interaction of Src and RACK1 Is

Enhanced by Activation of Protein Kinase C and Tyrosine Phosphorylation of RACK1. *J. Biol. Chem.* **276**, 20346–20356 (2001).

144. Rigas, A. C., Ozanne, D. M., Neal, D. E. & Robson, C. N. The Scaffolding Protein RACK1 Interacts with Androgen Receptor and Promotes Cross-talk through a Protein Kinase C Signaling Pathway. *J. Biol. Chem.* **278**, 46087–46093 (2003).

TABLES:

Table 1: Description of mains variables of the samples in association with type of diagnosis

		PVL *n* (%)	**OL** *n* (%)	**Control** *n* (%)	***p-value*[§]**
Substrate	saliva	7 (58.3)	8 (61.5)	5 (55.56)	0.961
	tissue	5 (41.7)	5 (38.5)	4 (44.4)	
Gender	men	1 (8.3)	9 (69.2)	2 (22.2)	0.004
	women	11 (91.7)	4 (30.8)	7 (77.8)	
Race	Caucasian	7 (58.3)	8 (61.5)	6 (66.7)	0.486
	Afro-American	0 (0)	5 (38.5)	1 (11.1)	
	Latin-American	5 (41.7)	0 (0)	2 (22.2)	
	Asian	0 (0)	0 (0)	0 (0)	
Clinical presentation	homogeneous	0 (0)	8 (61.5)	-	0.000
	verrucous	12 (100)	1 (7.7)	-	
	nodular	0 (0)	2 (15.4)	-	
	erythroleukoplakia	0 (0)	2 (15.4)	-	
Biopsy location	jugal mucosa	3 (25)	3 (23.1)	-	0.029
	tongue	3 (25)	8 (61.5)	-	
	mouth floor	0 (0)	0 (0)	-	
	palate	0 (0)	1 (7.7)	-	
	keratinized gingival tissue	3 (25)	1 (7.7)	-	
	alveolar ridge	2 (16.7)	0 (0)	-	
	labial mucosa	1 (8.3)	0 (0)	-	
	vestibule fundus	0 (0)	0 (0)	-	
	retromolar trigone	0 (0)	0 (0)	-	
Associated comorbidities	without pathology	7 (58.3)	7 (53.8)	6 (66.7)	0.171
	cardiovascular	1 (8.3)	2 (15.4)	0 (0)	
	endocrine[†]	1 (8.3)	2 (15.4)	2 (22.2)	
	psychiatric	0 (0)	2 (15.4)	0 (0)	
	degenerative[††]	1 (8.3)	0 (0)	0 (0)	
	Multiple[‡]	3 (25)	0 (0)	1 (11.1)	

Dysplasia	without dysplasia	3 (25)	4 (30.8)	9 (100)	0.331
	low grade	4 (33.3)	7 (53.8)	-	
	high grade	5 (41.7)	2 (15.4)	-	
Tobacco	Non smoker	10 (83.3)	4 (30.8)	6 (66.7)	0.044
	smoker	2 (16.7)	7 (53.8)	1 (11.1)	
	former smoker	0 (0)	2 (15.4)	2 (22.2)	
Alcohol consumption	no	10 (83.3)	12 (92.3)	7 (77.8)	0.621
	yes	2 (16.7)	1 (7.7)	2 (22.2)	

† *Endocrine: Diabetes and / or hypothyroidism.*
†† *Degenerative: Alzheimer's, multiple sclerosis, rheumatoid arthritis, Parkinson's.*
‡ *Multiple: Two or more of the above mentioned.*
§ *p-value correspond to chi^2 test*

Table 2. Primary antibodies used in IHC experiments to validate potential biomarkers.

Protein Target	Antibody	Clone	Manufacturer	Dilution
CALR	mouse anti-human calreticulin mAb	FMC 75	Abcam, Cambridge, UK	1:2000
GNB2L1	anti-human GNB2L1 Rabbit pAb	A0151	ABclonal, Woburn, USA	1:200
YWHAQ	anti-human YWHAQ Rabbit pAb	A2563	ABclonal, Woburn, USA	1:200

Table 3: Differential protein expression analysis between Tissue & Saliva

Protein name	Gene name	Tissue: FC PVL/OL	Tissue: FC PVL/control	Tissue: DE q-value	Protein name	Gene name	Saliva: FC PVL/OL	Saliva: FC PVL/control	Saliva: DE q-value
40S ribosomal protein SA	RPSA§	2.1	1.6	0.00	40S ribosomal protein SA	RPSA	-0.3	-0.7	0.20
Histone H4	HIST1H4A†	1.9	1.6	0.02	Histone H4	HIST1H4A	1.2	1.0	0.34
14-3-3 protein epsilon	YWHAE§	1.7	2.1	0.00	14-3-3 protein epsilon	YWHAE	-0.2	-0.3	0.43
Alpha-enolase	ENO1†	1.6	1.4	0.03	Alpha-enolase	ENO1	-0.5	-1.0	0.01
14-3-3 protein beta/alpha	YWHAB§	1.5	1.0	0.01	14-3-3 protein beta/alpha	YWHAB	0.1	-0.1	0.99
Serpin B5	SERPINB5§	1.4	1.4	0.00	Serpin B5	SERPINB5	-0.9	-1.2	0.01
Heat shock 70 kDa protein 1B	HSPA1B†	1.4	1.3	0.02	Heat shock 70 kDa protein 1B	HSPA1B	0.3	0.1	0.44
14-3-3 protein theta	YWHAQ§	1.3	1.4	0.02	14-3-3 protein theta	YWHAQ	-1.3	-2.1	0.06
Fructose-bisphosphate aldolase A	ALDOA†	1.3	0.9	0.01	Fructose-bisphosphate aldolase A	ALDOA	-0.5	-1.1	0.00
Heat shock cognate 71 kDa protein	HSPA8†	1.3	1.3	0.03	Heat shock cognate 71 kDa protein	HSPA8	-0.3	-0.3	0.26
Desmoglein-1	DSG1‡	1.1	1.6	0.04	Desmoglein-1	DSG1	-1.0	-1.7	0.00
Keratin, type II cytoskeletal 1	KRT1‡	1.1	1.5	0.04	Keratin, type II cytoskeletal 1	KRT1	-1.6	-0.6	0.04
Rab GDP dissociation inhibitor beta	GDI2§	1.0	1.2	0.00	Rab GDP dissociation inhibitor beta	GDI2	-0.6	-0.9	0.02
Glyceraldehyde -3-phosphate dehydrogenase	GAPDH†	1.0	0.4	0.01	Glyceraldehyde -3-phosphate dehydrogenase	GAPDH	-0.4	-1.4	0.07
ATP synthase subunit beta, mitochondrial	ATP5B‡	0.9	1.1	0.02	ATP synthase subunit beta, mitochondrial	ATP5F1B	-1.2	-0.6	0.11
Keratin, type I cytoskeletal 10	KRT10‡	0.8	1.4	0.04	Keratin, type I cytoskeletal 10	KRT10	-1.1	-0.9	0.03
14-3-3 protein zeta/delta	YWHAZ‡	0.5	0.9	0.03	14-3-3 protein zeta/delta	YWHAZ	-0.5	-0.7	0.08

Keratin, type I cytoskeletal 16	KRT16‡	0.3	2.2	0.01	Keratin, type I cytoskeletal 16	KRT16	-0.7	-0.6	0.09
Keratin, type I cytoskeletal 14	KRT14‡	0.1	1.2	0.02	Keratin, type I cytoskeletal 14	K1C14	-1.3	-1.8	0.03
Keratin, type II cytoskeletal 4	KRT4‡	-1.1	-1.9	0.04	Keratin, type II cytoskeletal 4	KRT4	-0.5	-0.2	0.35
Keratin, type I cytoskeletal 19	KRT19†	2.8	-0.2	0.01	Alpha-2-HS-glycoprotein	FETUA	0.0	-1.6	0.39
Suprabasin	SBSN†	1.8	1.4	0.03	Angiotensinogen	ANGT	0.0	-2.4	0.45
Receptor of activated protein C kinase 1	GNB2L1§	1.7	1.6	0.00	Hemopexin	HEMO	-0.1	-1.9	0.35
Calreticulin	CALR§	1.7	1.7	0.00	Hemoglobin subunit delta	HBD	-0.3	-1.4	0.23
Actin, alpha cardiac muscle 1	ACTC1§	1.7	1.9	0.01	Transforming protein RhoA	RHOA	-0.4	-2.2	0.04
60S ribosomal protein L22	RPL22§	1.5	0.9	0.00	Ceruloplasmin	CERU	-0.5	-1.8	0.23
Protein disulfide-isomerase A3	PDIA3†	1.4	0.9	0.01	6-phosphogluconate dehydrogenase, decarboxylating	6PGD	-0.7	-1.4	0.01
Desmocollin-3	DSC3§	1.4	1.6	0.01	Immunoglobulin kappa variable 1-5	KV105	-0.8	-1.6	0.27
Transitional endoplasmic reticulum ATPase	VCP§	1.3	1.0	0.01	Nucleoside diphosphate kinase B	NDKB	-1.0	-2.2	0.22
Tubulin beta chain	TUBB†	1.2	1.5	0.00	Adenosylhomocysteinase	SAHH	-1.0	-1.8	0.00
Keratin, type II cytoskeletal 78	KRT78†	1.2	0.1	0.04	Chloride intracellular channel protein 1	CLIC1	-1.2	-2.2	0.00
X-ray repair cross-	XRCC5§	1.1	1.5	0.00	Prostaglandin reductase 1	PTGR1	-1.3	-1.6	0.00

complementing protein 5									
60S ribosomal protein L4	RPL4§	1.0	0.8	0.00	Cell division control protein 42 homolog	CDC42	-1.3	-2.1	0.00
Neuroblast differentiation-associated protein AHNAK	AHNAK†	0.9	1.1	0.01	Cocaine esterase	EST2	-1.3	-1.7	0.00
Filamin-B	FLNB§	0.9	0.9	0.02	Cornulin	CRNN	-1.3	-1.6	0.06
Filamin-A	FLNA‡	0.9	1.4	0.01	Ras-related protein Rap-1A	RAP1A	-1.4	-1.6	0.00
Periplakin	PPL†	0.9	0.2	0.02	Transgelin-2	TAGL2	-1.4	-1.5	0.00
Catenin alpha-1	CTNNA1§	0.8	0.9	0.00	Interleukin-1 receptor antagonist protein	IL1RA	-1.4	-1.4	0.00
Keratin, type I cytoskeletal 17	KRT17‡	0.7	2.4	0.03	ERO 1-like protein alpha	ERO1A	-1.5	-1.8	0.00
Heat shock protein HSP 90-alpha	HSP90AA1‡	0.7	1.9	0.00	Beta-2-microglobulin	B2MG	-1.7	-1.4	0.00
Myosin-9	MYH9‡	0.6	1.7	0.02	Liver carboxylesterase 1	EST1	-1.7	-2.0	0.02
60S ribosomal protein L35a	RPL35A‡	0.6	1.5	0.00	Proteasome activator complex subunit 2	PSME2	-1.7	-2.3	0.00
Plectin	PLEC§	0.6	0.7	0.02	Cystatin-S	CYTS	-1.9	-1.7	0.10
Desmoplakin	DSP‡	0.6	0.9	0.01	BPI fold-containing family B member 2	BPIB2	-1.9	-1.7	0.02
Ubiquitin-like modifier-activating enzyme 1	UBA1†	0.5	0.4	0.04	Involucrin	INVO	-1.9	-2.8	0.25
Dermokine	DMKN‡	0.5	2.3	0.00	Zymogen granule protein	ZG16B	-2.0	-1.5	0.15

					16 homolog B				
Protein S100-A16	S100A16‡	-0.1	-1.6	0.00	Interleukin-36 alpha	IL36A	-2.0	-2.5	0.00
Ras-related protein Rab-11A	RAB11A‡	-0.2	0.5	0.01	Ly6/PLAUR domain-containing protein 3	LYPD3	-2.0	-2.5	0.01
					Carboxypeptidase E	CBPE	-2.1	-1.8	0.01
					Heme-binding protein 2	HEBP2	-2.1	-1.8	0.00
					Antileukoproteinase	SLPI	-2.2	-2.1	0.11
					Small proline-rich protein 3	SPRR3	-2.2	-1.5	0.07
					BPI fold-containing family A member 1	BPIA1	-2.4	-2.5	0.06
					IgGFc-binding protein	FCGBP	-2.4	-1.8	0.03
					Putative cytochrome P450 2D7	CP2D7	-2.8	-3.0	0.05
					BPI fold-containing family B member 1	BPIB1	-2.8	-2.7	0.14
					Lipocalin-1	LCN1	-2.9	-2.8	0.15
					Lysozyme C	LYSC	-4.0	-1.7	0.19
					Mesothelin	MSLN	-4.5	-3.2	0.01

PVL: proliferative verrucous leukoplakia, OL: oral leukoplakia, FC: foldchange values for each specific comparison between groups, DE: differential analysis between PVL group against OL and control groups considering FDR=0.01

§ proteins that show difference between PVL against OL and control groups at the same time in tissue samples

† proteins that only show difference between PVL and OL groups in tissue samples

‡ proteins that only show difference between PVL and control groups in tissue samples

Difference on tissue samples in KRT4, KRT78, KRT1, KRT10, UBA1 proteins were observed based only in foldchange values due the post hoc cannot see where is the difference between groups

Cut-off for FC values was ±1.2 for tissue and saliva samples

The upper part of the table is delimited by the thick black line and shows the same proteins identified in tissue and saliva, the lower part of the table shows different proteins identified in tissue and saliva
Proteins had been organized by the FC PVL/OL values in decreasing order with reference in the tissue samples to the upper part of the table and lower part was organized in decreasing order independently for each substratum

Table 4: PVL Multiple comparisons

Proteins	Mean expression §	SD*	Grade of Dysplasia†		Difference of means	95% Confidence Interval		p-value‡
						Lower limit	Upper limit	
RPSA	0.069	1.018	Low	Without	0.35	-0.24	0.94	0.34
			High	Without	2.13	1.56	2.70	0.00
				Low	1.78	1.32	2.24	0.00
YWHAQ[ƒ]	0.167	0.952	Low	Without	-	-	-	-
			High	Without	-	-0.71	1.05	0.00
				Low	-	-0.71	1.05	0.00
YWHAB	0.055	1.029	Low	Without	0.56	-0.62	1.73	0.57
			High	Without	2.14	1.01	3.28	0.00
				Low	1.59	0.68	2.50	0.00
SERPINB5	0.107	0.973	Low	Without	0.54	0.04	1.03	0.04
			High	Without	2.15	1.67	2.63	0.00
				Low	1.62	1.23	2.00	0.00
GDI2	0.052	1.032	Low	Without	0.24	-0.93	1.40	1.00
			High	Without	1.97	0.85	3.10	0.00
				Low	1.74	0.84	2.64	0.00
YWHAE	0.065	1.022	Low	Without	0.32	-0.53	1.17	0.87
			High	Without	2.08	1.25	2.90	0.00
				Low	1.76	1.10	2.41	0.00
ALDOA	0.074	1.014	Low	Without	0.74	-0.70	2.18	0.47
			High	Without	2.11	0.72	3.50	0.01
				Low	1.37	0.25	2.48	0.02
GAPDH	0.083	1.005	Low	Without	0.88	-0.59	2.34	0.33
			High	Without	2.14	0.72	3.55	0.01
				Low	1.26	0.12	2.39	0.03
ENO1	0.073	1.015	Low	Without	0.71	-1.08	2.49	0.80
			High	Without	1.94	0.22	3.67	0.03
				Low	1.23	-0.15	2.62	0.08
HSPA1B	0.050	1.033	Low	Without	1.01	-0.79	2.81	0.39
			High	Without	2.10	0.36	3.83	0.02
				Low	1.08	-0.31	2.48	0.14
HSPA8	0.056	1.029	Low	Without	0.70	-0.55	1.94	0.39
			High	Without	2.19	0.99	3.39	0.00
				Low	1.49	0.53	2.46	0.00
HIST1H4A	0.008	1.048	Low	Without	0.30	-0.88	1.49	1.00
			High	Without	2.04	0.89	3.19	0.00
				Low	1.74	0.82	2.66	0.00
KRT14	0.058	1.028	Low	Without	0.27	-0.23	0.76	0.43
			High	Without	2.11	1.63	2.59	0.00
				Low	1.84	1.46	2.23	0.00
KRT16	0.062	1.024	Low	Without	0.08	-0.55	0.71	1.00
			High	Without	1.97	1.36	2.58	0.00
				Low	1.89	1.40	2.38	0.00
YWHAZ	0.062	1.025	Low	Without	0.34	-0.60	1.28	0.92
			High	Without	2.07	1.17	2.98	0.00
				Low	1.73	1.01	2.46	0.00

DSG1	0.056	1.029	Low	Without	-0.01	-0.74	0.73	1.00
			High	Without	1.91	1.19	2.62	0.00
				Low	1.91	1.34	2.48	0.00
KRT1	0.046	1.035	Low	Without	0.42	-0.76	1.60	0.95
			High	Without	2.08	0.94	3.22	0.00
				Low	1.66	0.75	2.58	0.00
KRT10	0.053	1.031	Low	Without	0.45	-0.57	1.47	0.66
			High	Without	2.13	1.14	3.12	0.00
				Low	1.68	0.89	2.47	0.00
KRT4	0.049	1.033	Low	Without	0.08	-0.64	0.79	1.00
			High	Without	1.97	1.28	2.66	0.00
				Low	1.90	1.34	2.45	0.00
ATP5B	0.047	1.049	Low	Without	0.47	-0.58	1.52	0.58
			High	Without	2.27	1.17	3.38	0.00
				Low	1.81	0.88	2.73	0.00

§ Mean expression correspond to the area below the curve for the m/z detection of every protein in each sample, with log2(x) transformation and z-score normalization by protein

** Standard deviation*

† Comparison across the different grades of dysplasia

‡ α=0.05 for ANOVA and post hoc Bonferroni, note that in all significant differences the "high grade dysplasia" always had a greater amount of each protein detected

ƒ In the YWHAQ protein was not possible stablish the parameters in the post hoc because 5 samples with high grade were compared with 1 without and 1 low grade of dysplasia , notwithstanding, ANOVA test showed significance difference

FIGURES:

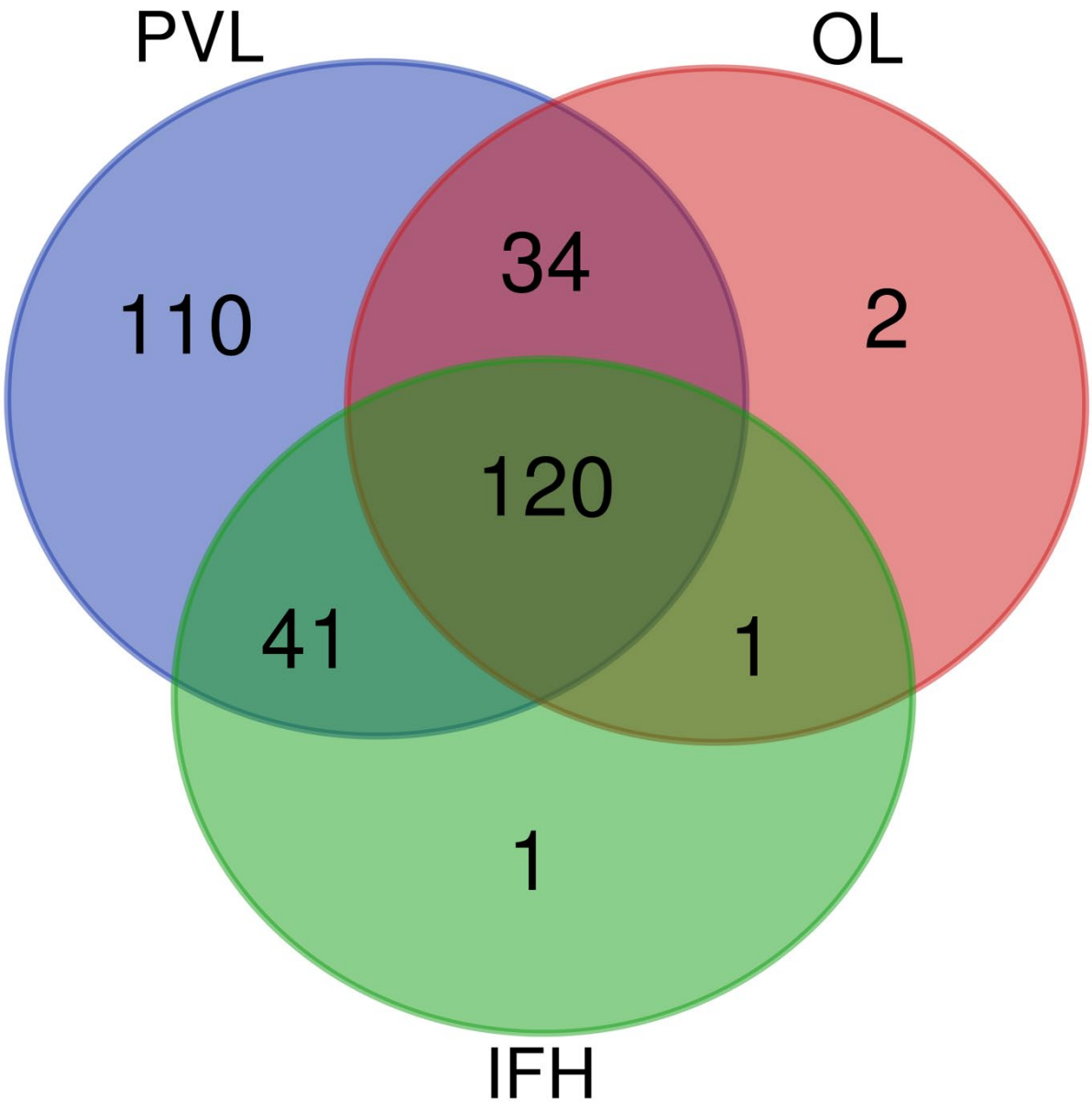


Figure 1. Venn diagram of proteins identified on tissue samples

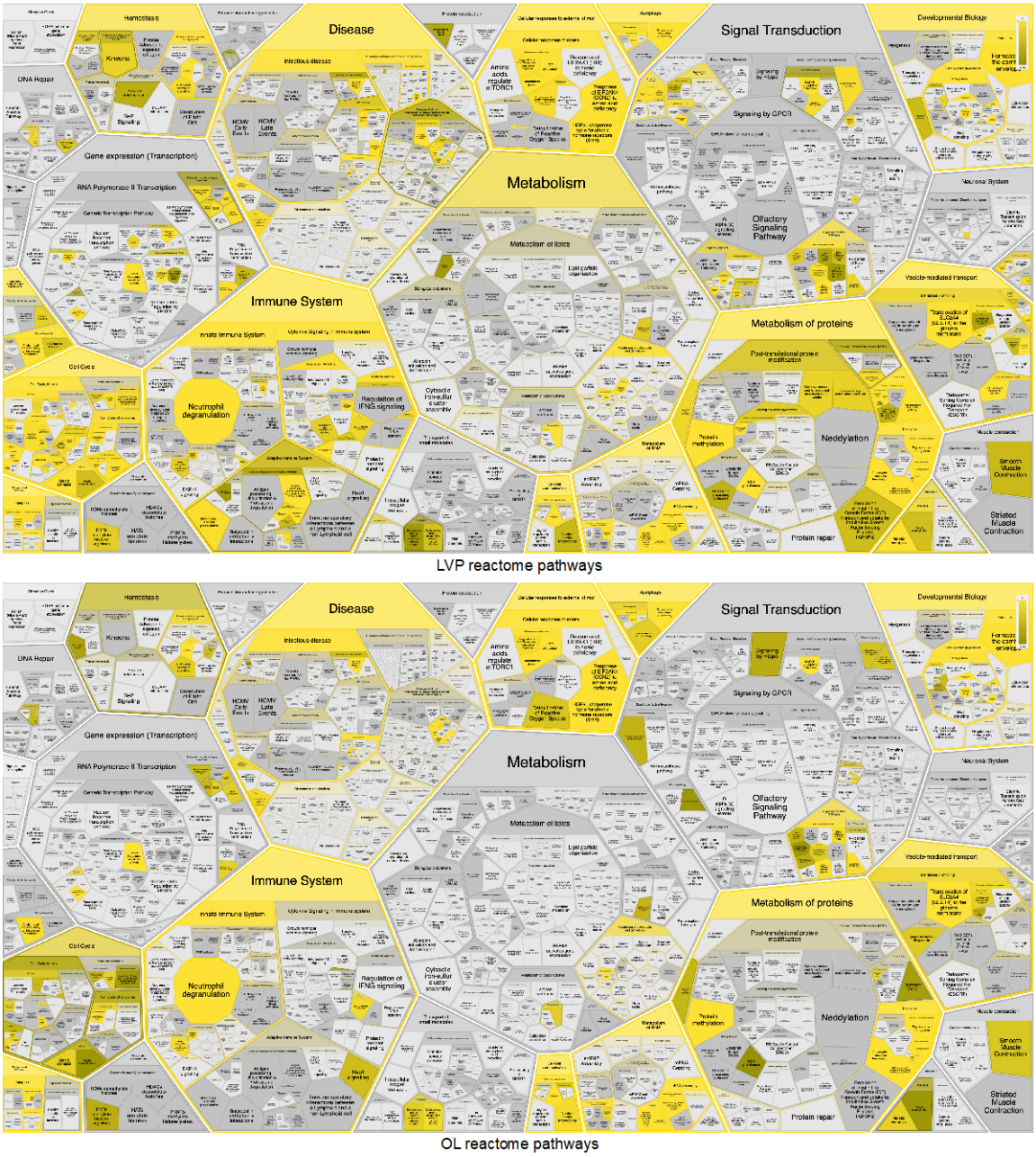


Figure 2. Reactome pathways of PVL and OL samples from tissue samples

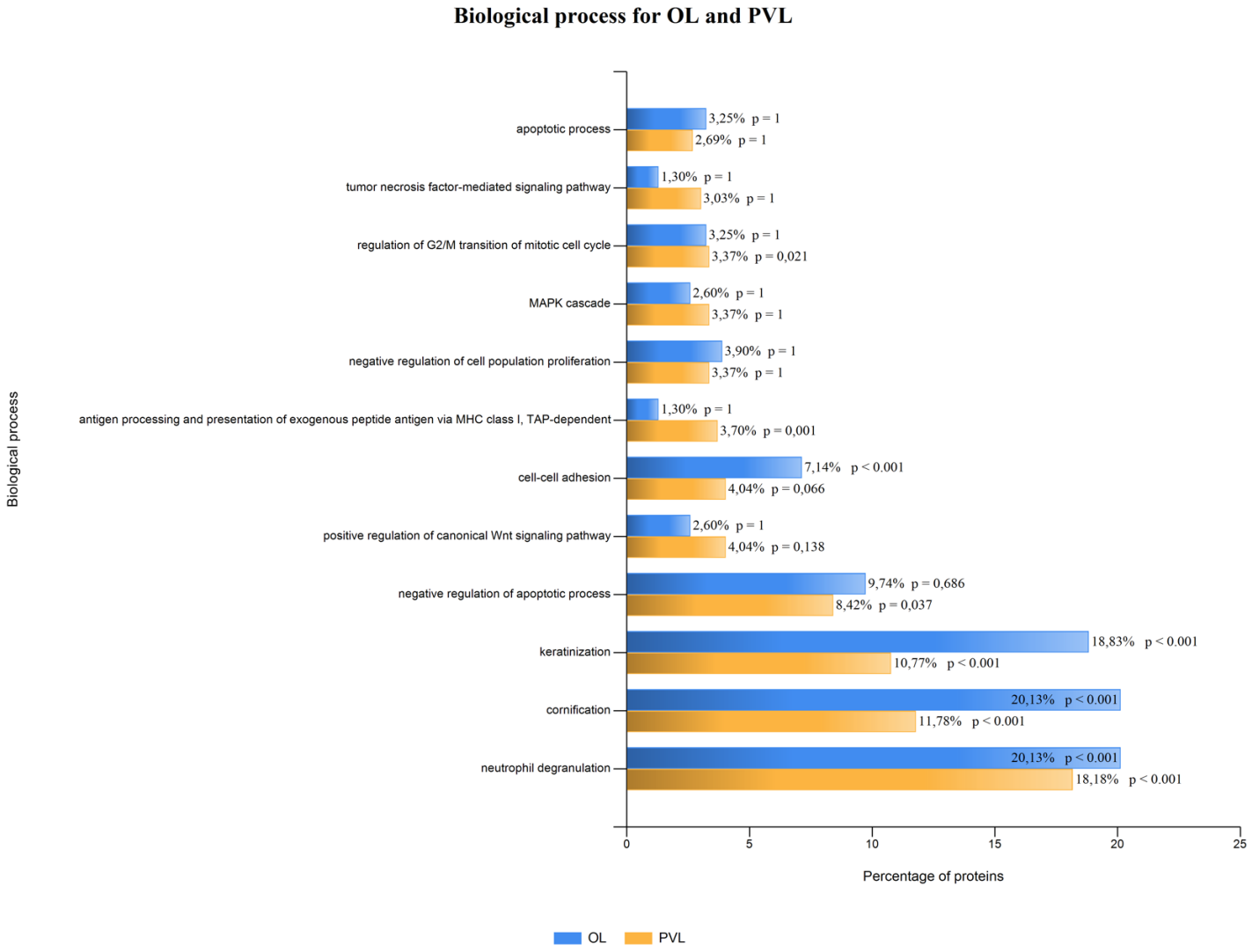


Figure 3. Biological processes assessed by FunRich tool in PLV and OL tissue proteins

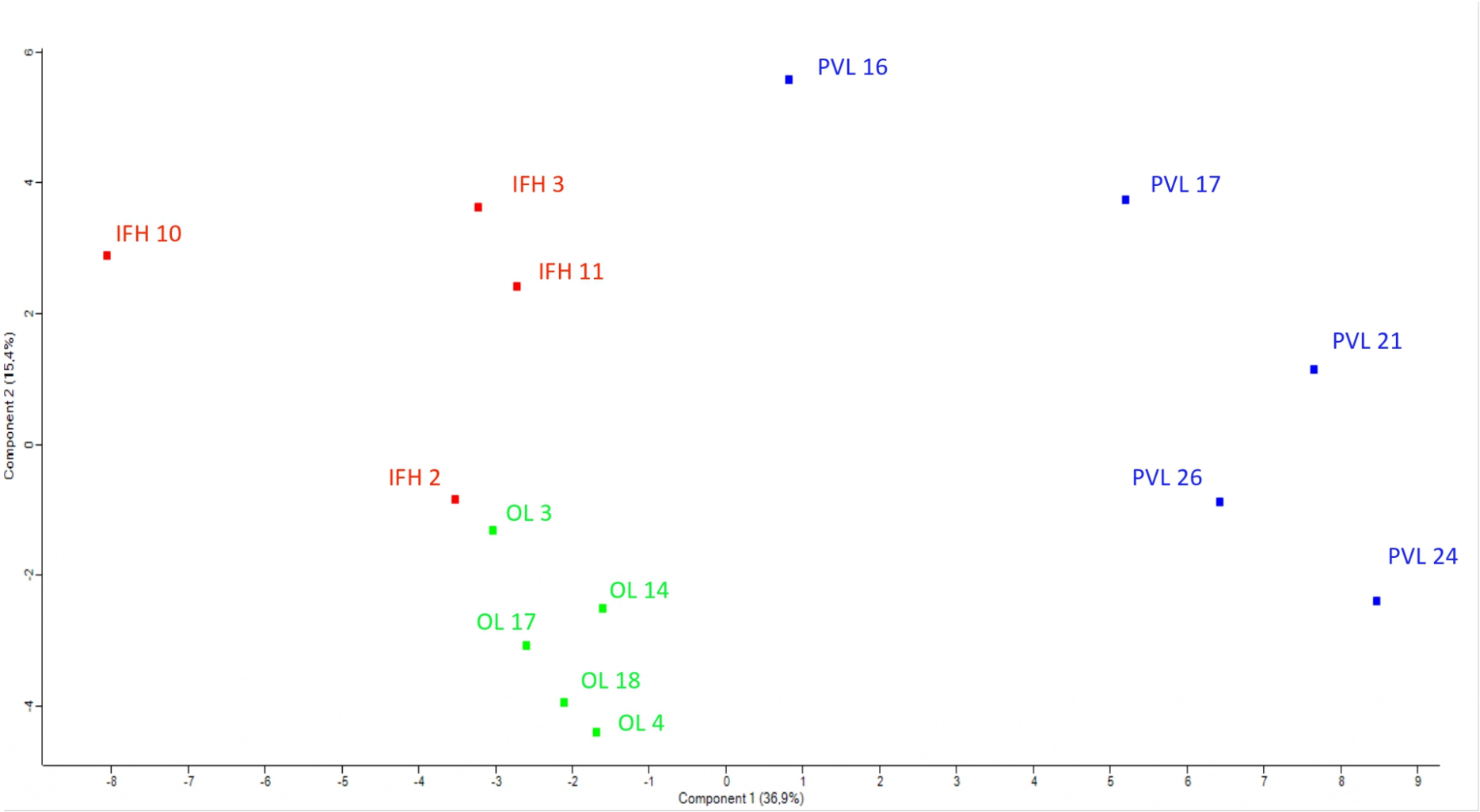


Figure 4. Principal component analysis (PCA) of tissue samples

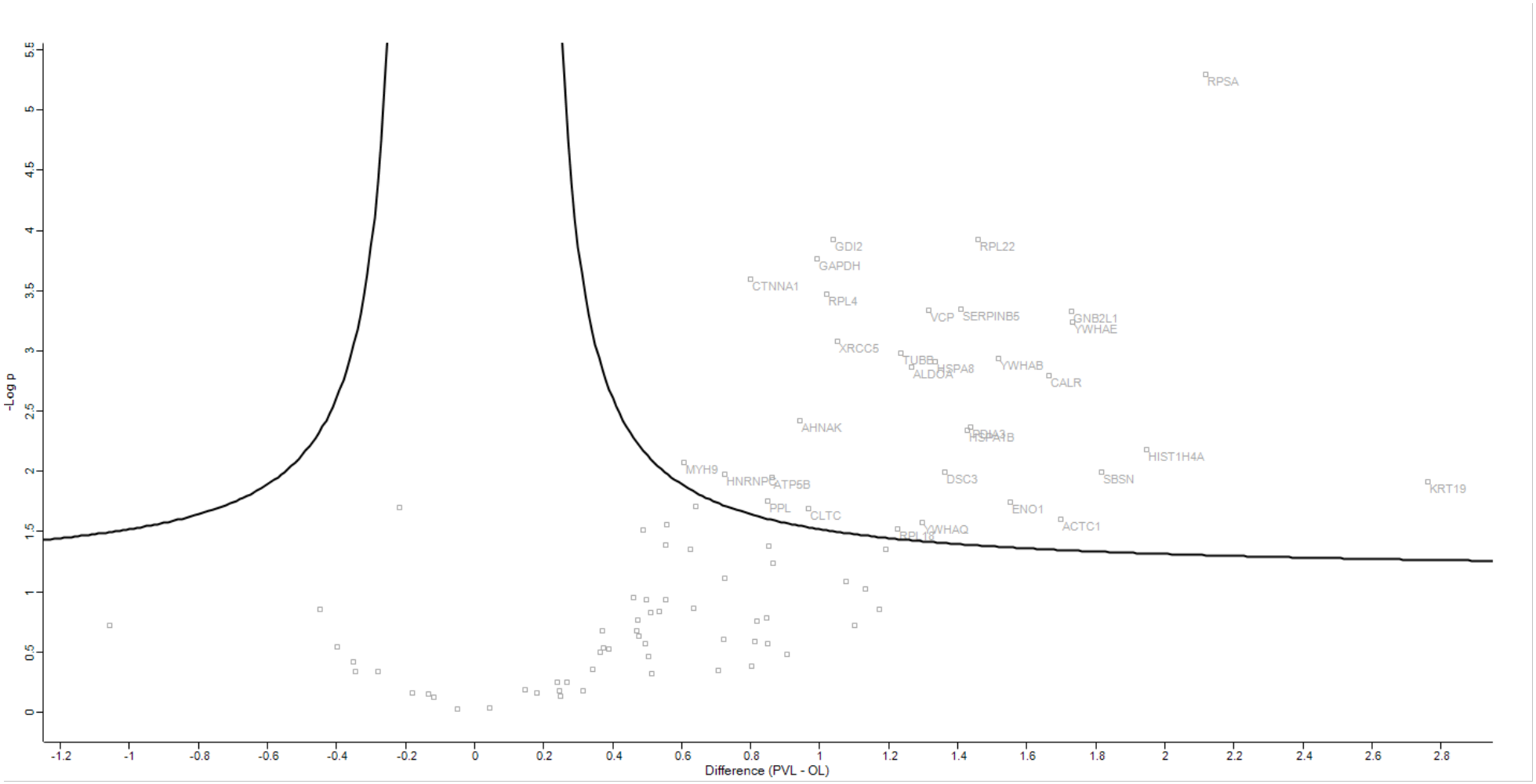


Figure 5. Volcano plot; Foldchange assessment of PLV/OL

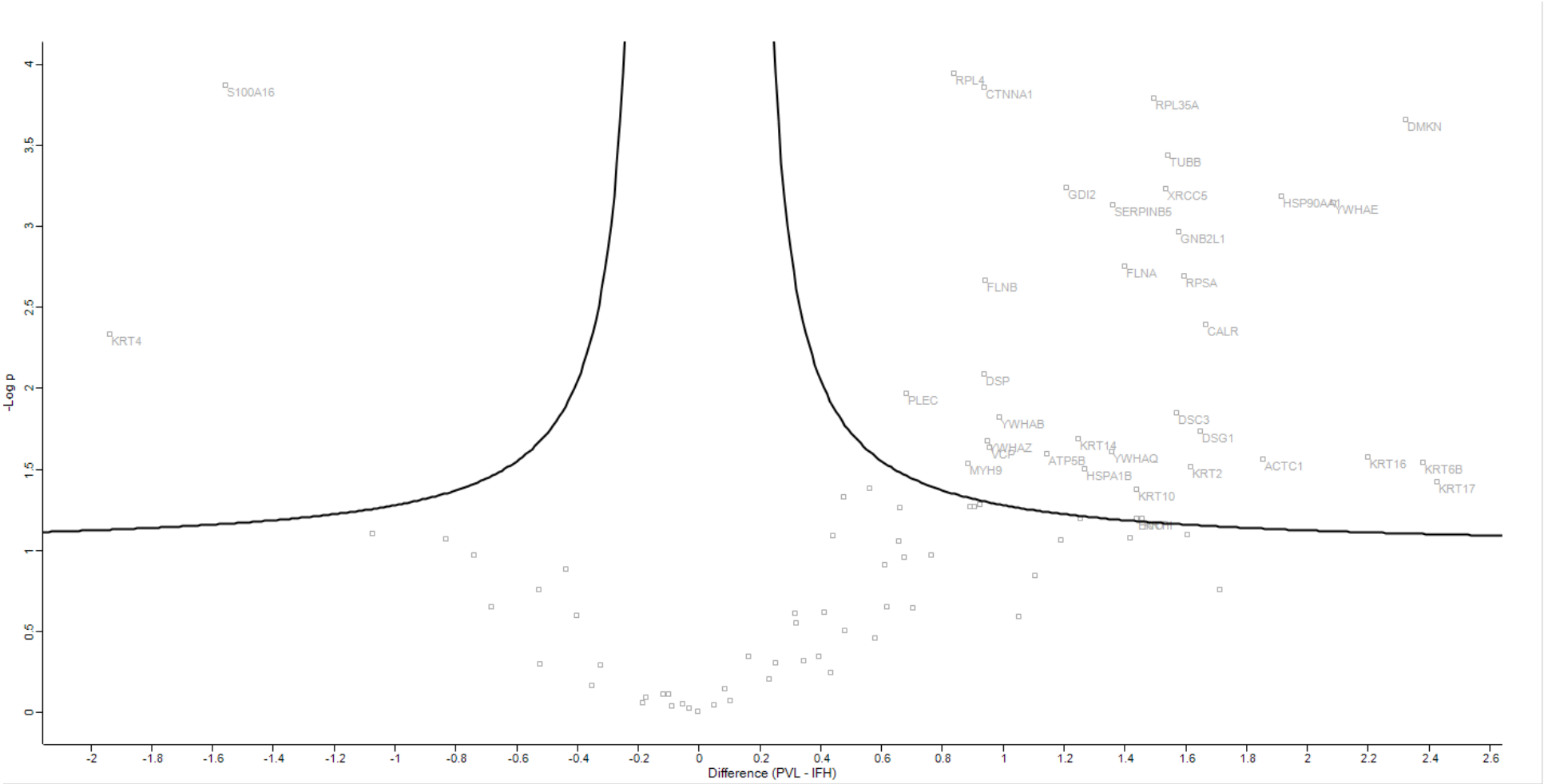


Figure 6. Volcano plot; Foldchange assessment of PLV/IFH

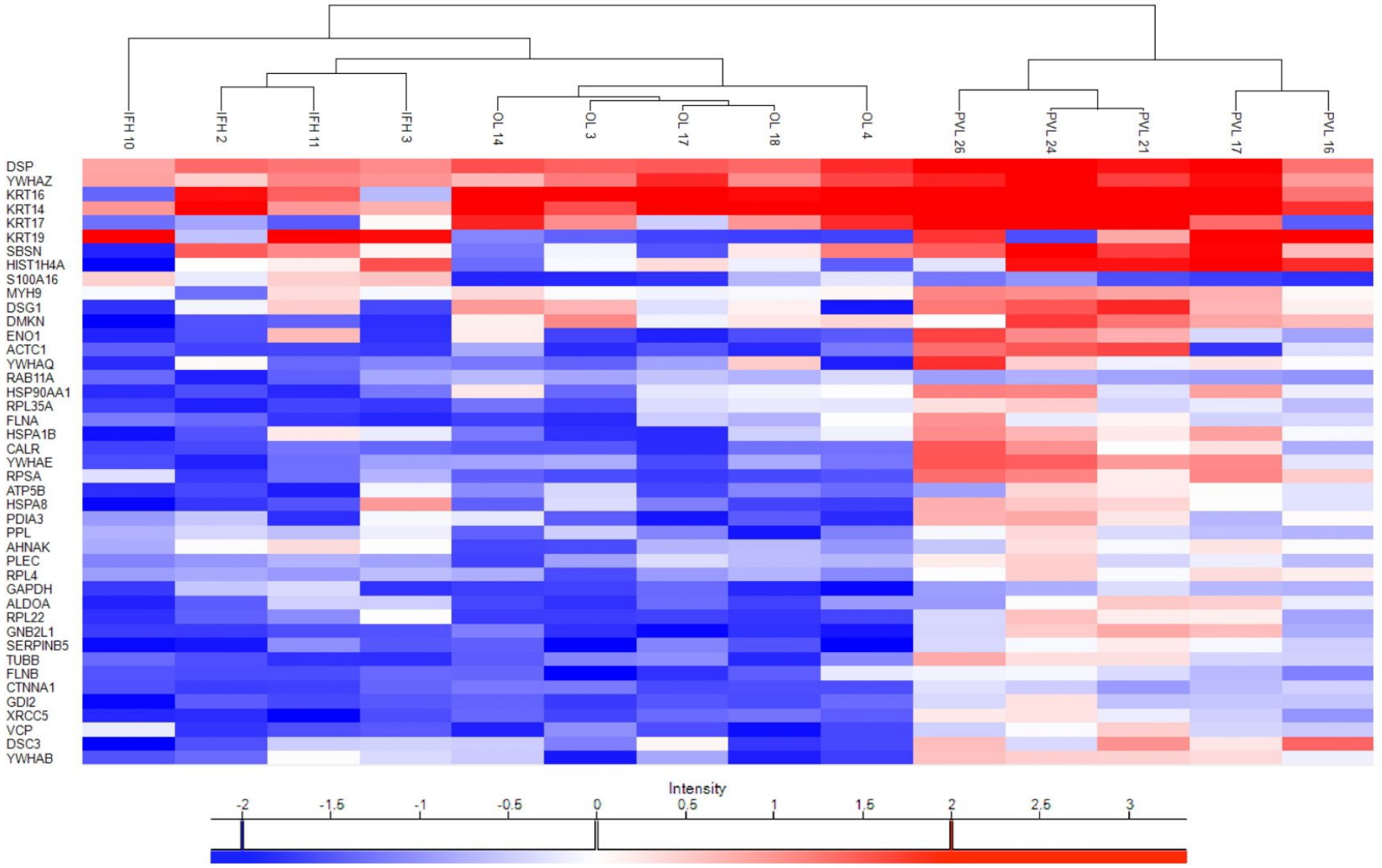


Figure 7. Heatmap of potential biomarkers in tissue samples

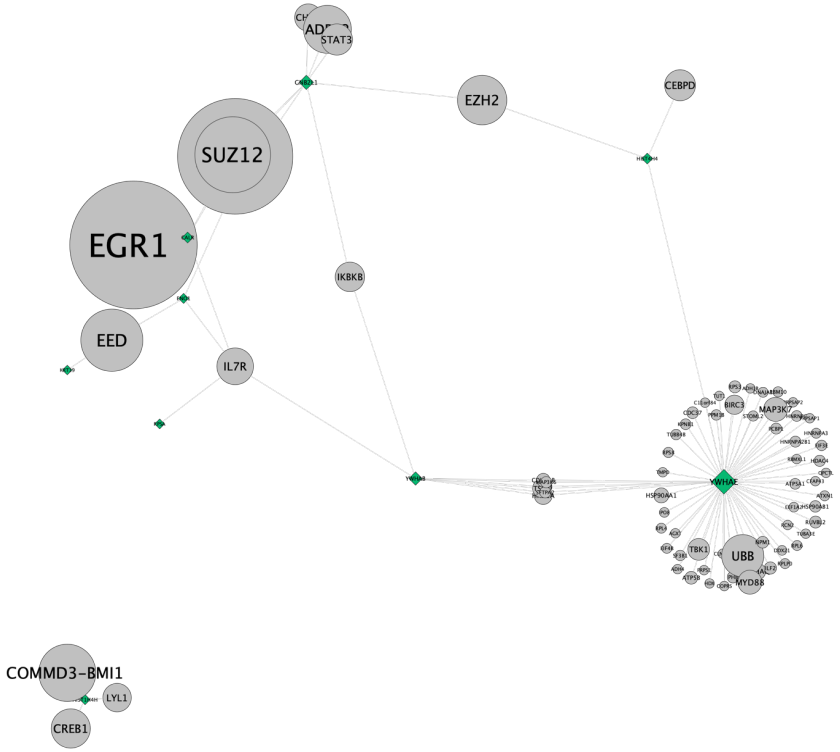


Figure 8. Bioinformatic functional network analysis by CHAT on Cytoscape

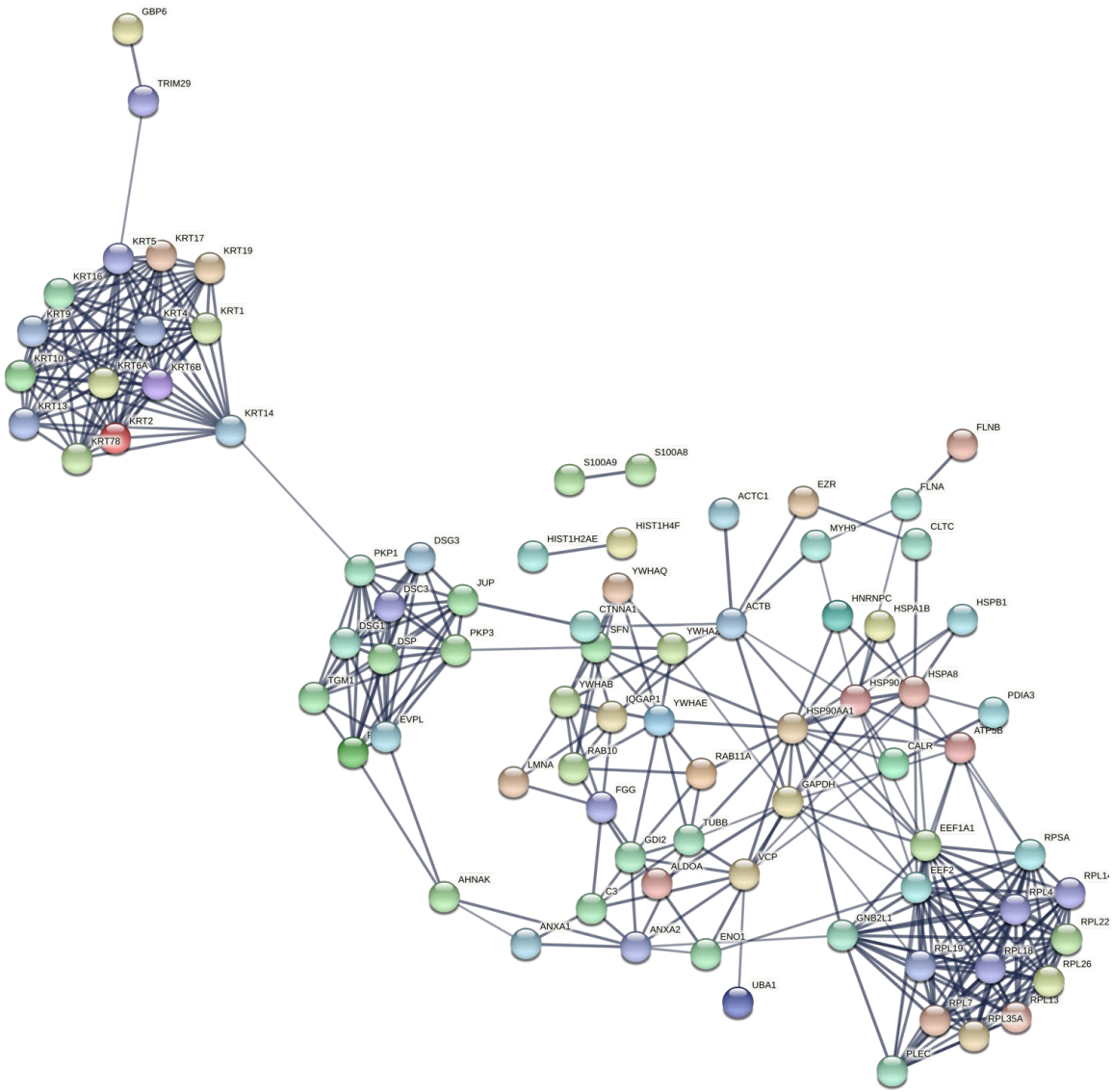


Figure 9. Bioinformatic functional network analysis by String

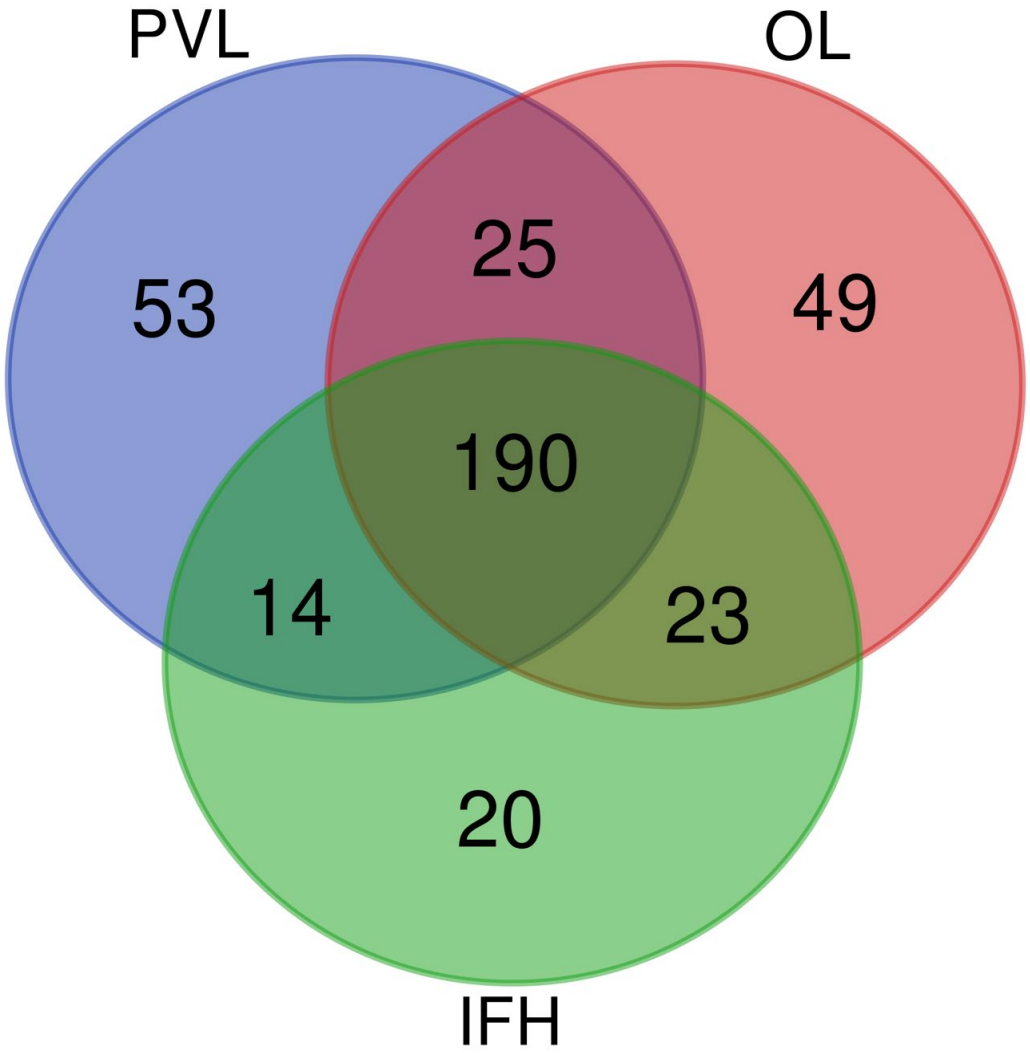
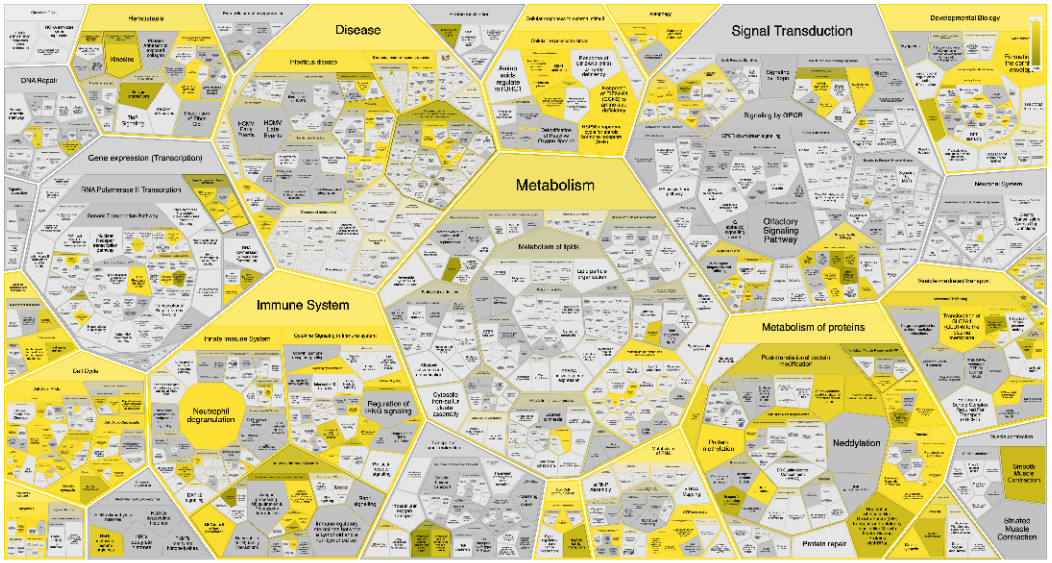
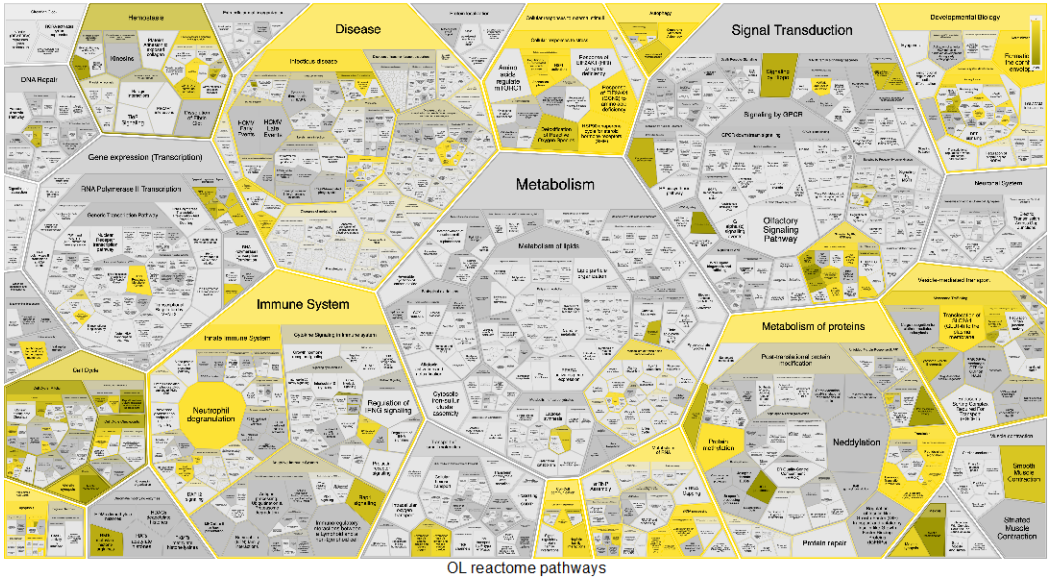


Figure 10. Venn diagram of proteins identified on saliva samples



LVP reactome pathways



OL reactome pathways

Figure 11. Reactome pathways of PVL and OL samples from saliva samples

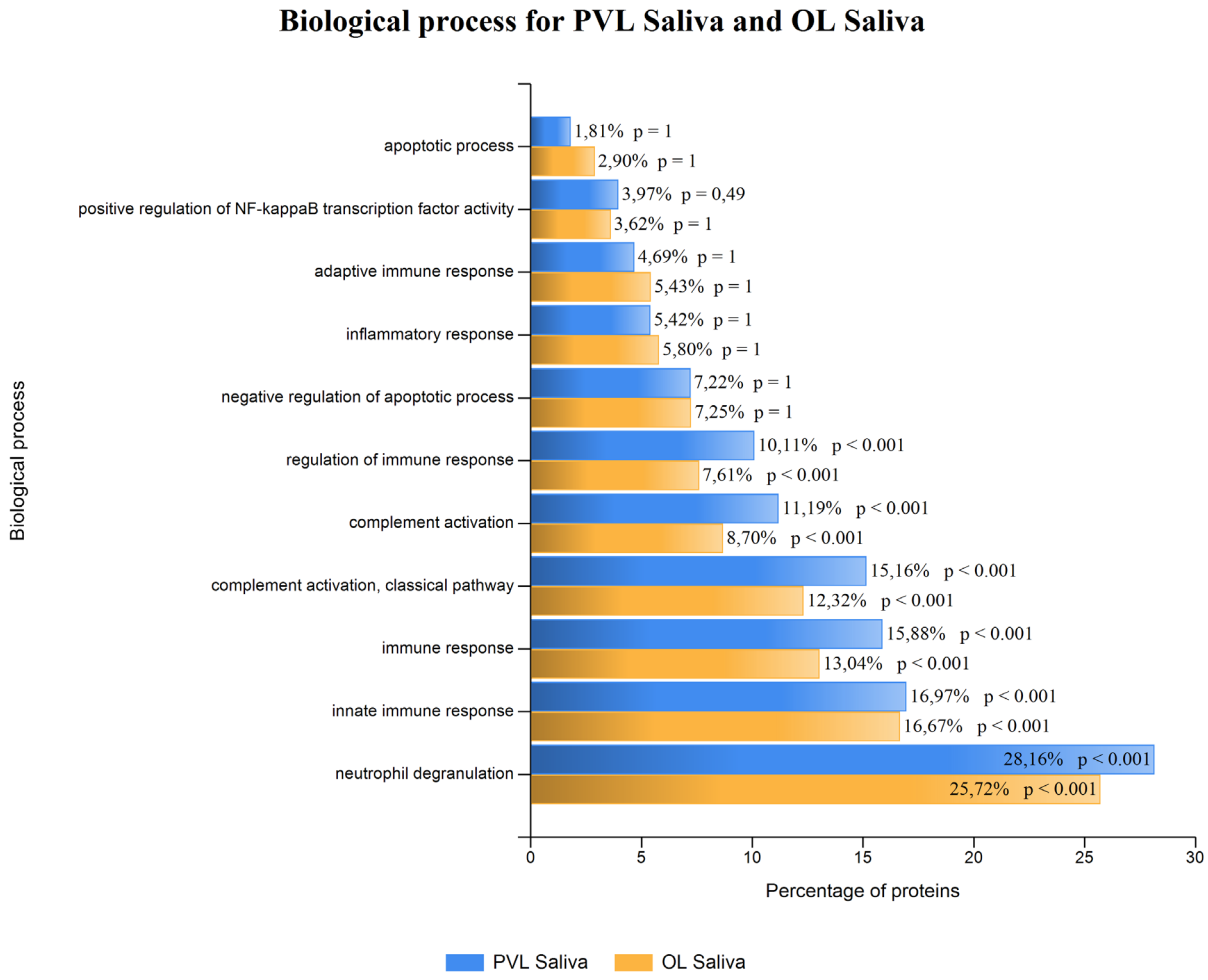


Figure 12. Biological processes assessed by FunRich tool in PLV and OL saliva proteins

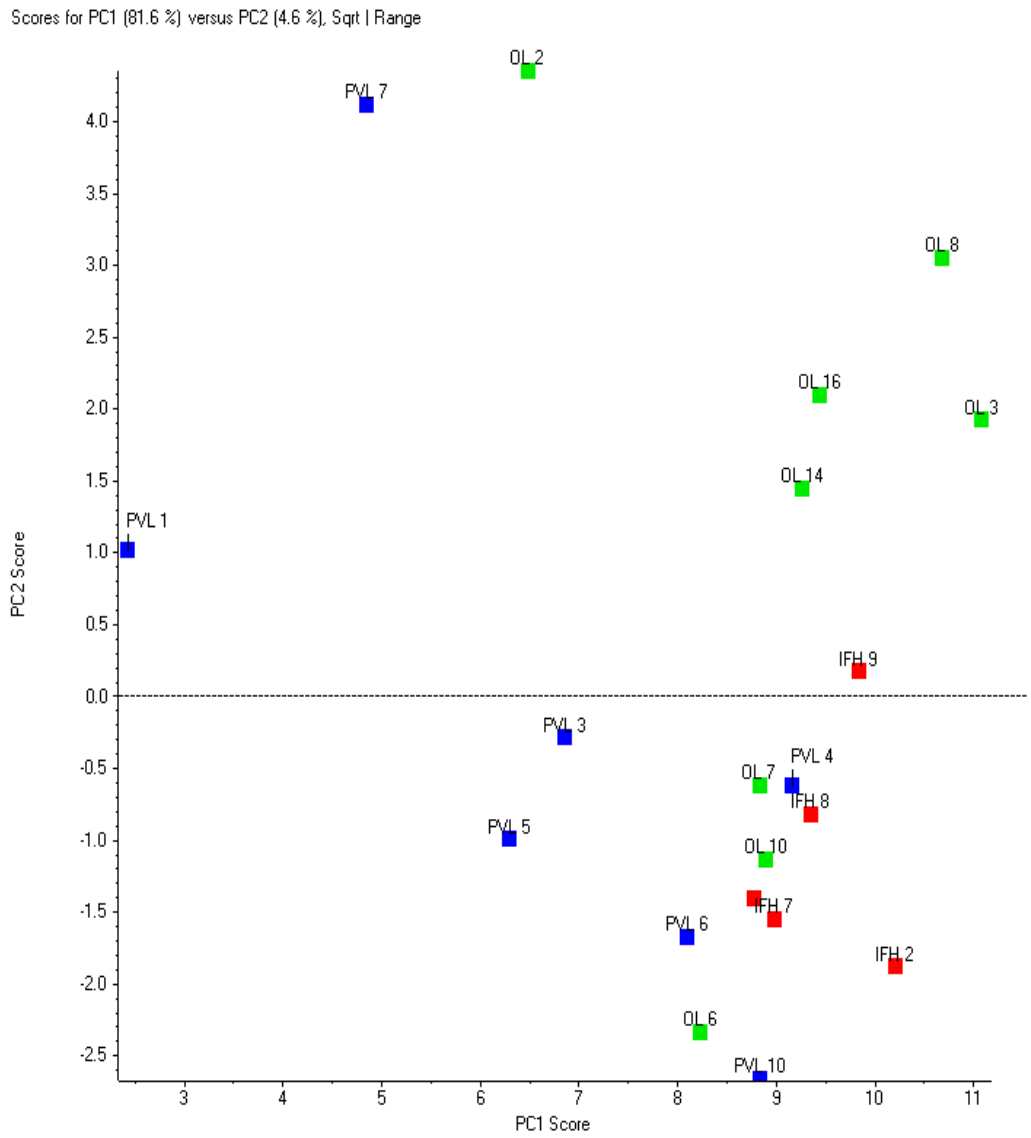


Figure 13. Principal component analysis (PCA) of saliva samples

SUPPLEMENTARY TABLES:

Table 1: Clinical data base of samples used in LC-MS/MS

Codification	Diagnosis	Substrate	Sex	Age	Biopsy location	Associated comorbidities	Dysplasia degree	Previous carcinoma	Clinical evolution in the period studied
GLVP_S_001	PVL	Saliva	Female	85	Keratinized gingiva	Without	Without	Carcinoma in situ	Stable
GLVP_S_003	PVL	Saliva	Female	69	Keratinized gingiva	Without	Low	Not	Stable
GLVP_S_004	PVL	Saliva	Female	51	Jugal mucosa	Without	Without	Not	Healing
GLVP_S_005	PVL	Saliva	Female	72	Keratinized gingiva	Without	Low	Not	Stable
GLVP_S_006	PVL	Saliva	Female	94	Jugal mucosa	Without	-	Not	Stable
GLVP_S_007	PVL	Saliva	Male	83	Tongue	Degenerative	Low	Not	Healing
GLVP_S_010	PVL	Saliva	Female	87	Jugal mucosa	Without	Low	Not	Healing
GL_S_002	OL	Saliva	Male	61	Tongue	Without	High	Not	Healing
GL_S_003	OL	Saliva	Female	62	Tongue	Psychiatric	Without	Not	Stable
GL_S_006	OL	Saliva	Male	86	Jugal mucosa	Without	Low	Not	Healing
GL_S 007	OL	Saliva	Female	66	Palate	Cardiovascular	Without	Not	Stable
GL_S_008	OL	Saliva	Male	28	Tongue	Without	Low	Not	Recurrence
GL_S_010	OL	Saliva	Male	44	Tongue	Without	Low	Not	Stable
GL_S_014	OL	Saliva	Male	82	Tongue	Psychiatric	Low	Not	New disease
GL_S_016	OL	Saliva	Male	58	Jugal mucosa	Without	Without	Not	Stable
GC_S_002	Control	Saliva	Female	37	-	Without	Without	Not	Healing
GC_S_006	Control	Saliva	Female	36	-	Without	Without	Not	Healing
GC_S_007	Control	Saliva	Female	70	-	Without	Without	Not	Healing
GC_S_008	Control	Saliva	Male	52	-	Without	Without	Not	Healing
GC_S_009	Control	Saliva	Male	47	-	Without	Without	Not	Healing
PVL 16	PVL	Tissue	Female	73	Alveolar rim	Without	High	Not	Recurrence
PVL 17	PVL	Tissue	Female	67	Tongue	Cardiovascular	High	Not	Recurrence

PVL 21	PVL	Tissue	Female	52	Tongue	Dyslipidemia, osteoporosis, depression	High	Carcinoma in situ	Oncological event
PVL 24	PVL	Tissue	Female	55	Lip (internal mucosa)	Osteoporosis, depression	High	Not	Oncological event
PVL 26	PVL	Tissue	Female	71	Alveolar rim	Hypertension, hypothyroidism, acute myocardial infarction, had lung cancer	High	Not	Oncological event
OL 3	OL	Tissue	Male	50	Jugal mucosa	-	Low	Not	Stable
OL 4	OL	Tissue	Male	51	Keratinized gingiva	Cardiovascular	Low	Not	Stable
OL 14	OL	Tissue	Female	60	Tongue	Endocrine; diabetes and / or hypothyroidism	Without	Not	Stable
OL17	OL	Tissue	Female	47	Tongue	Endocrine; diabetes and / or hypothyroidism	High	Carcinoma in situ	Oncological event
OL 18	OL	Tissue	Male	55	Tongue	Without	Low	Not	Recurrence
IFH 2	Control	Tissue	Female	49	Lip (internal mucosa)	Endocrine; diabetes and / or hypothyroidism	Without	Not	Healing
IFH 3	Control	Tissue	Female	67	Alveolar rim	Endocrine; diabetes and / or hypothyroidism	Without	Not	Healing
IFH 10	Control	Tissue	Female	53	Vestibule fundus	Without	Without	Not	Healing
IFH 11	Control	Tissue	Female	76	Alveolar rim	Aasthma, vitiligo, diverticulitis	Without	Not	Healing

Table 2: Differential protein expression analysis in Saliva between PLV and OL groups

Protein	Gene	p-value	FC PVL/OL
Serpin B5	SERPINB5[§]	0.04	-4.6
Fructe-bisphphate aldolase A	ALDOA[†]	0.05	-4.4
14-3-3 protein zeta/delta	YWHAZ[†]	0.14	-2.8
14-3-3 protein epsilon	YWHAE[§]	0.58	-0.8
14-3-3 protein beta/alpha	YWHAB[§]	0.89	-0.2
Desmoglein-1	DSG1[†]	0.01	-6.2
Keratin, type I cytkeletal 14	KRT14[†]	0.03	-5.2
14-3-3 protein theta	YWHAQ[§]	0.10	-3.3
Alpha-enolase	ENO1[†]	0.10	-3.3
Rab GDP dissociation inhibitor beta	GDI2[§]	0.11	-3.2
Keratin, type I cytkeletal 16	KRT16[†]	0.12	-3.1
Glyceraldehyde-3-phphate dehydrogenase	GAPDH[†]	0.18	-2.5
Heat shock cognate 71 kDa protein	HSPA8[†]	0.37	-1.4
40S ribomal protein SA	RPSA[§]	0.42	-1.2
Histone H4	HIST1H4A[†]	0.42	-1.2
Heat shock 70 kDa protein 1B	HSPA1B[†]	0.45	-1.1
14-3-3 protein sigma	SFN	0.02	-5.6
Prtaglandin reductase 1	PTGR1	0.00	-10.9
Heme-binding protein 2	HEBP2	0.00	-10.4
Interleukin-36 alpha	IL36A	0.00	-9.9
Cocaine esterase	CES2	0.00	-9.8
Interleukin-1 receptor antagonist protein	IL1RN	0.00	-8.3
Cofilin-1	CFL1	0.00	-8.1
Cell division control protein 42 homolog	CDC42	0.00	-8.1
Mesothelin	MSLN	0.00	-8.0
Nucleobindin-1	NUCB1	0.00	-8.0
Kallikrein-11	KLK11	0.00	-7.9
Annexin A1	ANXA1	0.00	-7.9
Protein CutA	CUTA	0.00	-7.8
ERO1-like protein alpha	ERO1A	0.00	-7.8
Glutathione S-transferase P	GSTP1	0.00	-7.7
Plastin-3	PLS3	0.01	-7.5
Adenylhomocysteinase	AHCY	0.01	-7.5
Proteasome activator complex subunit 1	PSME1	0.01	-7.4
Beta-2-microglobulin	B2M	0.01	-7.3
Proteasome activator complex subunit 2	PSME2	0.01	-7.3
Transgelin-2	TAGLN2	0.01	-7.0
Carboxypeptidase E	C	0.01	-7.0
Desmocollin-2	DSC2	0.01	-6.9
Protein FAM3D	FAM3D	0.01	-6.7
Nucleobindin-2	NUCB2	0.01	-6.4
Nucleotide exchange factor SIL1	SIL1	0.01	-6.3
Thioredoxin	TXN	0.02	-5.9

Platelet-activating factor acetylhydrolase IB subunit beta	PAFAH1B2	0.02	-5.9
Ly6/PLAUR domain-containing protein 3	LYPD3	0.02	-5.8
Keratin, type I cytkeletal 13	KRT13	0.02	-5.8
Peroxiredoxin-1	PRDX1	0.02	-5.7
BPI fold-containing family B member 2	BPIFB2	0.02	-5.7
Galectin-3-binding protein	LGALS3BP	0.02	-5.4
Keratin, type II cytkeletal 6C	KRT6C	0.02	-5.4
Cysteine-rich secretory protein 3	CRISP3	0.02	-5.3
Serpin B3	SERPINB3	0.03	-5.3
Malate dehydrogenase, mitochondrial	MDH2	0.03	-5.3
Ras-related protein Rap-1A	RAP1A	0.03	-5.3
Chloride intracellular channel protein 1	CLIC1	0.03	-5.3
Liver carboxylesterase 1	CES1	0.03	-5.3
Peptidyl-prolyl cis-trans isomerase B	PPIB	0.03	-5.2
Zinc-alpha-2-glycoprotein	AZGP1	0.03	-5.2
Trypsin-1	PRSS1	0.03	-5.2
Putative ubiquitin-conjugating enzyme E2 N-like	UBE2NL	0.03	-5.2
IgGFc-binding protein	FCGBP	0.03	-5.1
Apolipoprotein A-II	APOA2	0.03	-4.9
Metalloproteinase inhibitor 1	TIMP1	0.03	-4.9
Phphatidylethanolamine-binding protein 1	BP1	0.03	-4.8
Fatty acid-binding protein, epidermal	FABP5	0.04	-4.7
Keratin, type I cytkeletal 10	KRT10	0.04	-4.5
GTP-binding nuclear protein Ran	RAN	0.04	-4.5
Cathepsin Z	CTSZ	0.04	-4.5
Keratin, type II cytkeletal 2 epidermal	KRT2	0.05	-4.4
Serpin B13	SERPINB13	0.05	-4.4
Moesin	MSN	0.05	-4.4
Cystatin-S	CST4	0.05	-4.4

PVL: proliferative verrucous leukoplakia, OL: oral leukoplakia, FC: foldchange values for each specific comparison between groups.
§ proteins that show difference between PVL against OL and CG at the same time in tissue samples
† proteins that only show difference between PVL and OL groups in tissue samples

Table 3: Differential protein expression analysis in Saliva between PLV and CG groups

Group	Gene	p-value	FC PVL/CG
Fructe-bisphphate aldolase A	ALDOA†	0.00	-1.1
Desmoglein-1	DSG1†	0.00	-1.7
Alpha-enolase	ENO1†	0.00	-1.0
Serpin B5	SERPINB5§	0.02	-1.2
Keratin, type I cytkeletal 14	KRT14†	0.05	-1.8
Rab GDP dissociation inhibitor beta	GDI2§	0.05	-0.9
Keratin, type I cytkeletal 10	KRT10†	0.05	-0.9
14-3-3 protein theta	YWHAQ§	0.06	-2.1
Keratin, type II cytkeletal 1	KRT1†	0.07	-0.9
40S ribomal protein SA	RPSA§	0.10	-0.7
14-3-3 protein zeta/delta	YWHAZ†	0.11	-0.7
Keratin, type I cytkeletal 16	KRT16†	0.20	-0.6
Heat shock cognate 71 kDa protein	HSPA8†	0.37	-0.3
14-3-3 protein epsilon	YWHAE§	0.44	-0.3
Complement C3	C3†	0.50	-0.4
Histone H4	HIST1H4A†	0.57	1.0
Keratin, type II cytkeletal 4	KRT4†	0.68	-0.2
Heat shock 70 kDa protein 1B	HSPA1B†	0.77	0.1
14-3-3 protein beta/alpha	YWHAB†	0.86	-0.1
14-3-3 protein sigma	SFN	0.17	-0.7
Transforming protein RhoA	RHOA	0.00	-2.2
Cocaine esterase	CES2	0.00	-1.7
Chloride intracellular channel protein 1	CLIC1	0.00	-2.2
Adenylhomocysteinase	AHCY	0.00	-1.8
Interleukin-36 alpha	IL36A	0.00	-2.5
Proteasome activator complex subunit 2	PSME2	0.00	-2.3
6-phphogluconate dehydrogenase, decarboxylating	PGD	0.00	-1.4
Cathepsin B	CTSB	0.00	-1.3
ERO1-like protein alpha	ERO1A	0.00	-1.8
Cell division control protein 42 homolog	CDC42	0.00	-2.1
Aldehyde dehydrogenase, dimeric NADP-preferring	ALDH3A1	0.00	-1.3
Transmembrane protease serine 11D	TMPRSS11D	0.00	-1.1

Desmocollin-2	DSC2	0.00	-1.2
Heme-binding protein 2	HEBP2	0.00	-1.8
Peroxiredoxin-6	PRDX6	0.00	-1.0
Immunoglobulin heavy constant gamma 2	IGHG2	0.00	-1.4
Cofilin-1	CFL1	0.00	-1.1
Glutathione S-transferase P	GSTP1	0.00	-1.2
Cellular retinoic acid-binding protein 2	CRABP2	0.00	-1.4
Cornulin	CRNN	0.00	-1.6
Interleukin-1 receptor antagonist protein	IL1RN	0.01	-1.4
Malate dehydrogenase, mitochondrial	MDH2	0.01	-1.2
Prtaglandin reductase 1	PTGR1	0.01	-1.6
Putative cytochrome P450 2D7	CYP2D7	0.01	-3.0
NPC intracellular cholesterol transporter 2	NPC2	0.01	-1.0
Transgelin-2	TAGLN2	0.01	-1.5
Ras-related protein Rap-1A	RAP1A	0.01	-1.6
Ly6/PLAUR domain-containing protein 3	LYPD3	0.01	-2.5
BPI fold-containing family B member 1	BPIFB1	0.01	-2.7
Mucin-5B	MUC5B	0.01	-2.3
Lipocalin-1	LCN1	0.02	-2.8
Ras-related C3 botulinum toxin substrate 2	RAC2	0.02	-1.2
Nucleobindin-1	NUCB1	0.02	-1.7
L-lactate dehydrogenase B chain	LDHB	0.02	-1.3
Peroxiredoxin-1	PRDX1	0.02	-0.7
Glyceraldehyde-3-phphate dehydrogenase	GAPDH	0.02	-1.4
Ras-related protein Rab-10	RAB10	0.02	-0.9
L-lactate dehydrogenase A chain	LDHA	0.02	-1.0
Beta-2-microglobulin	B2M	0.02	-1.4
Annexin A1	ANXA1	0.02	-1.0
BPI fold-containing family A member 1	BPIFA1	0.02	-2.5
Carboxypeptidase E	C	0.02	-1.8
Antileukoproteinase	SLPI	0.02	-2.1
Poly(U)-specific endoribonuclease	ENDOU	0.02	-1.0
Alpha-2-macroglobulin-like protein 1	A2ML1	0.03	-1.2
Chitotriidase-1	CHIT1	0.03	2.2

Kallikrein-6	KLK6	0.03	1.1
Liver carboxylesterase 1	CES1	0.03	-2.0
Moesin	MSN	0.04	-0.9
Platelet-activating factor acetylhydrolase IB subunit beta	PAFAH1B2	0.04	-0.9
GTP-binding nuclear protein Ran	RAN	0.04	-1.3
Small proline-rich protein 3	SPRR3	0.04	-1.5
Coronin-1A	CORO1A	0.04	-1.2
Beta-enolase	ENO3	0.04	-1.1
BPI fold-containing family B member 2	BPIFB2	0.05	-1.7

PVL: proliferative verrucous leukoplakia, CG: control group, OL: oral leukoplakia, FC: foldchange values for each specific comparison between groups.
§ proteins that show difference between PVL against OL and CG at the same time in tissue samples
† proteins that only show difference between PVL and OL groups in tissue samples

Table 4: Descriptive statistics Marking Intensity of Immunohistochemistry assay

	CALR			YWHAQ			GNB2L		
Descriptive	**OL Grupo 1**	**PVL Grupo 2**	**IFH Grupo 3**	**OL Grupo 1**	**PVL Grupo 2**	**IFH Grupo 3**	**OL Grupo 1**	**PVL Grupo 2**	**IFH Grupo 3**
Minimum	5.64	9.02	4.28	9.49	0.07	7.59	0.17	0.36	6.77
Maximum	157.33	228.70	208.35	210.38	250.83	157.88	185.21	223.47	148.96
Amplitude Total	151.69	219.68	204.07	200.89	250.76	150.29	185.04	223.11	142.19
Median	64.42	96.23	88.68	125.44	122.21	92.14	108.58	86.24	28.24
First Quartile (25%)	37.16	67.91	57.12	75.67	69.21	36.64	18.43	37.21	18.07
Third Quartile (75%)	105.91	121.78	123.00	164.10	162.60	122.26	148.23	140.81	66.53
Interquartile Deviation	68.75	53.88	65.88	88.44	93.39	85.62	129.80	103.60	48.46
Arithmetic Mean	70.82	96.32	93.60	112.97	117.65	83.37	90.31	93.09	45.99
Variance	1720.53	1930.81	2392.38	3697.59	3978.74	3532.81	4473.49	3869.69	1982.47
Standard Deviation	41.48	43.94	48.91	60.81	63.08	59.44	66.88	62.21	44.53
Error Standard	6.33	4.10	8.65	11.70	6.65	24.27	13.12	6.42	14.08
Coefficient of Variation	0.59	0.46	0.52	0.54	0.54	0.71	0.74	0.67	0.97
Asymmetry (g1)	0.13	0.36	0.37	-0.30	0.03	-0.14	-0.17	0.20	1.60
Kurtosis (g2)	-1.06	0.31	-0.23	-1.02	-0.75	-1.87	-1.44	-0.96	2.36
Outlier 1	-169.09	-93.72	-140.51	-189.64	-210.97	-220.22	-370.98	-273.60	-127.30
Outlier 2	312.16	283.41	320.62	429.41	442.78	379.12	537.64	451.62	211.89

No outliers observed

Table 5: Pillai's trace for multivariate analyses

Effect	Value	F	Hypothesis df	Sig.	Partial Eta Squared	Observed Power
Biomarker	0.189	6.483	6.000	0.000	0.095	0.999

Biomarker: CALR, YWHAQ, GNB2L; F: F statistic; Hypothesis: number of dependent variables; df; degrees of freedom.

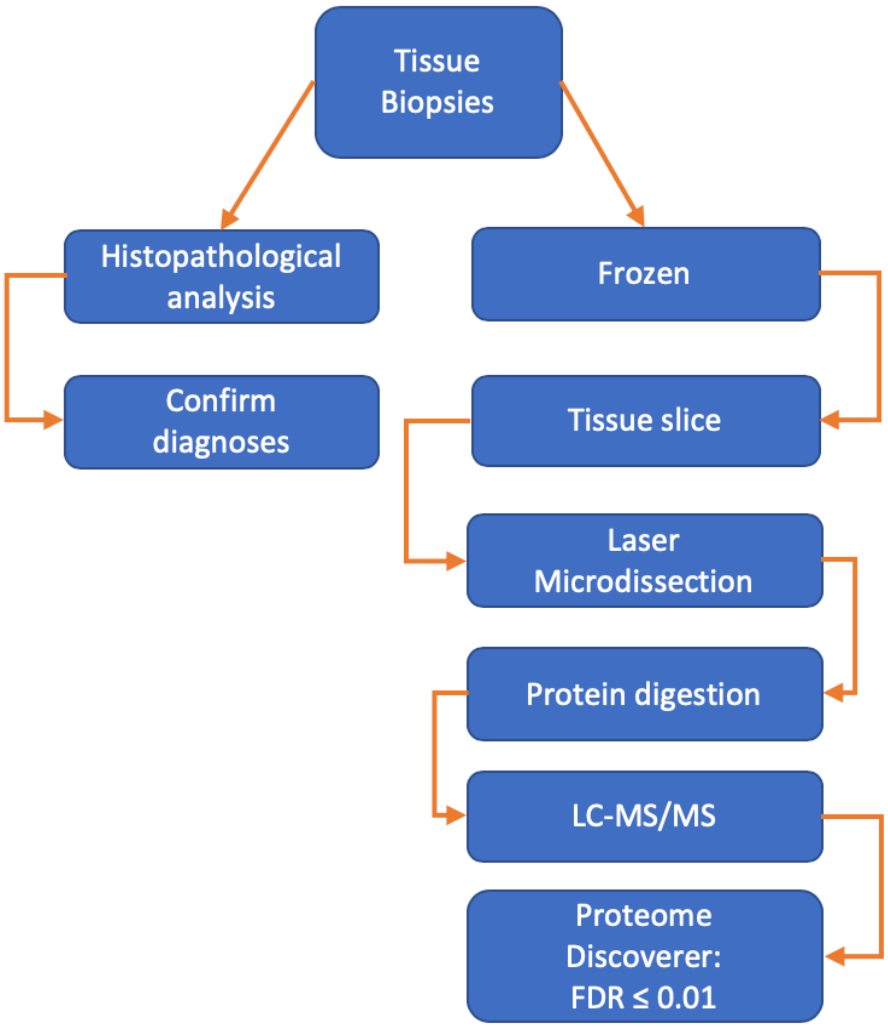
Table 6: Multiple Comparisons post hoc Games-Howell

Dependent Variable			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
CARL	OL	IFH	-22.78	10.71	0.09	-48.52	2.96
	PVL	OL	25.50*	7.54	0.00	7.49	43.50
		IFH	2.71	9.57	0.96	-20.46	25.89
YWHAQ	OL	IFH	29.61*	8.43	0.00	9.39	49.82
	PVL	OL	4.68	8.96	0.86	-16.68	26.04
		IFH	34.29*	6.69	0.00	18.40	50.17
GNB2L	OL	IFH	44.31*	8.94	0.00	22.85	65.77
	PVL	OL	2.78	9.45	0.95	-19.79	25.35
		IFH	47.10*	6.74	0.00	31.10	63.09

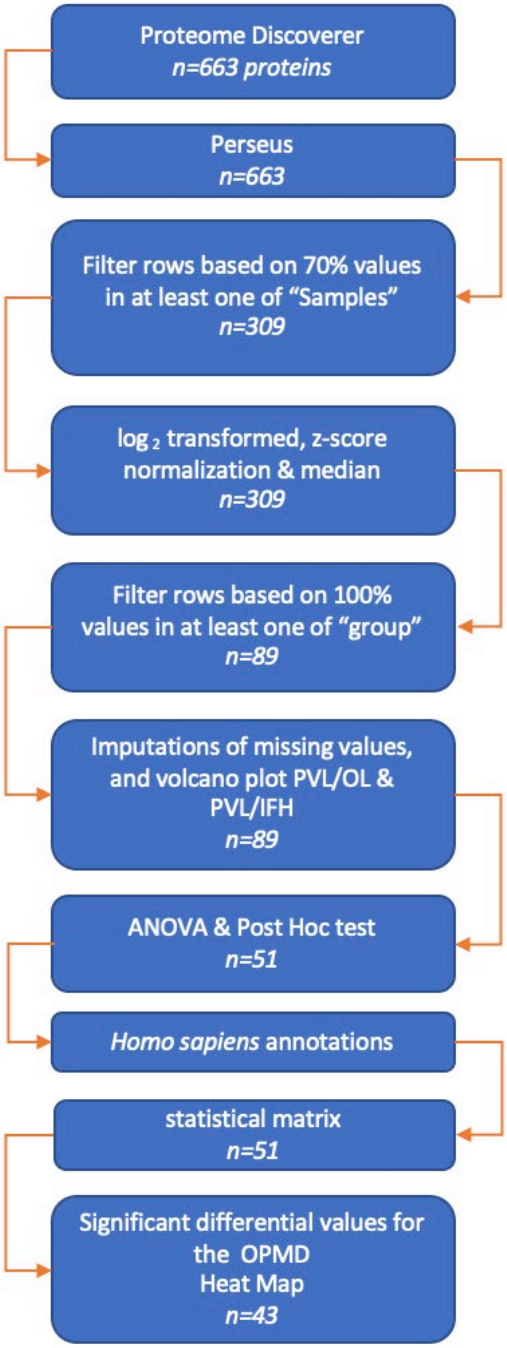
Coparissons based on observed means.
The error term is Mean Square(Error) = 2617,972.
**The mean difference is significant at the 0.05 level.*

SUPPLEMENTARY FIGURES:

Supplementary 1. Flow chart 1



Supplementary 2. Flow chart 2



Supplementary 3. Pathway Analysis Report 1

The following table shows the 25 most relevant pathways sorted by p-value.

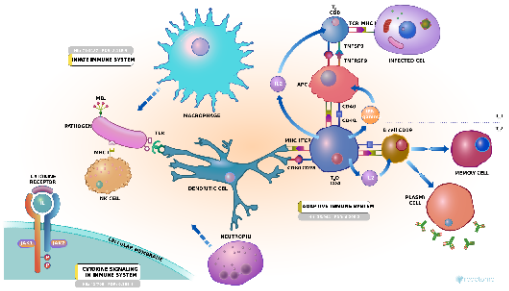
Pathway name	Entities				Reactions	
	found	ratio	p-value	FDR*	found	ratio
Eukaryotic Translation Elongation	31 / 95	0.008	1.11e-16	3.77e-15	9 / 9	6.94e-04
Peptide chain elongation	29 / 90	0.008	1.11e-16	3.77e-15	5 / 5	3.86e-04
Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	29 / 96	0.008	1.11e-16	3.77e-15	1 / 1	7.71e-05
Formation of a pool of free 40S subunits	28 / 102	0.009	1.11e-16	3.77e-15	2 / 2	1.54e-04
L13a-mediated translational silencing of Ceruloplasmin expression	30 / 112	0.01	1.11e-16	3.77e-15	3 / 3	2.31e-04
Viral mRNA Translation	27 / 101	0.009	1.11e-16	3.77e-15	2 / 2	1.54e-04
SRP-dependent cotranslational protein targeting to membrane	30 / 113	0.01	1.11e-16	3.77e-15	5 / 5	3.86e-04
GTP hydrolysis and joining of the 60S ribosomal subunit	30 / 113	0.01	1.11e-16	3.77e-15	3 / 3	2.31e-04
Nonsense-Mediated Decay (NMD)	29 / 117	0.01	1.11e-16	3.77e-15	6 / 6	4.63e-04
Nonsense Mediated Decay (NMD) enhanced by the Exon Junction Complex (EJC)	29 / 117	0.01	1.11e-16	3.77e-15	5 / 5	3.86e-04
Neutrophil degranulation	55 / 480	0.042	1.11e-16	3.77e-15	10 / 10	7.71e-04
Eukaryotic Translation Initiation	30 / 120	0.011	1.11e-16	3.77e-15	16 / 21	0.002
Cap-dependent Translation Initiation	30 / 120	0.011	1.11e-16	3.77e-15	13 / 18	0.001
Keratinization	35 / 217	0.019	1.11e-16	3.77e-15	23 / 34	0.003
Eukaryotic Translation Termination	27 / 94	0.008	1.11e-16	3.77e-15	3 / 5	3.86e-04
Formation of the cornified envelope	35 / 129	0.011	1.11e-16	3.77e-15	16 / 27	0.002
Selenocysteine synthesis	27 / 94	0.008	1.11e-16	3.77e-15	2 / 7	5.40e-04
Response of EIF2AK4 (GCN2) to amino acid deficiency	27 / 102	0.009	1.11e-16	3.77e-15	4 / 16	0.001
Cellular responses to stress	65 / 565	0.049	1.11e-16	3.77e-15	49 / 227	0.018
Cellular responses to external stimuli	65 / 579	0.051	1.11e-16	3.77e-15	49 / 258	0.02
Developmental Biology	94 / 1,100	0.096	1.11e-16	3.77e-15	101 / 537	0.041
Influenza Infection	32 / 172	0.015	1.11e-16	3.77e-15	10 / 59	0.005
Regulation of expression of SLITs and ROBOs	35 / 172	0.015	1.11e-16	3.77e-15	3 / 20	0.002
Selenoamino acid metabolism	27 / 118	0.01	1.11e-16	3.77e-15	2 / 33	0.003

Supplementary 4. Pathway Analysis Report 2

The following table shows the 25 most relevant pathways sorted by p-value.

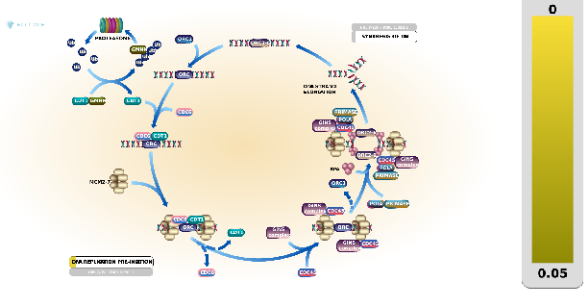
Pathway name	Entities				Reactions	
	found	ratio	p-value	FDR*	found	ratio
Peptide chain elongation	23 / 90	0.008	1.11e-16	3.33e-15	5 / 5	3.86e-04
Eukaryotic Translation Elongation	24 / 95	0.008	1.11e-16	3.33e-15	9 / 9	6.94e-04
Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	22 / 96	0.008	1.11e-16	3.33e-15	1 / 1	7.71e-05
Viral mRNA Translation	21 / 101	0.009	1.11e-16	3.33e-15	2 / 2	1.54e-04
Formation of a pool of free 40S subunits	21 / 102	0.009	1.11e-16	3.33e-15	2 / 2	1.54e-04
GTP hydrolysis and joining of the 60S ribosomal subunit	22 / 113	0.01	1.11e-16	3.33e-15	3 / 3	2.31e-04
Nonsense-Mediated Decay (NMD)	22 / 117	0.01	1.11e-16	3.33e-15	6 / 6	4.63e-04
Nonsense Mediated Decay (NMD) enhanced by the Exon Junction Complex (EJC)	22 / 117	0.01	1.11e-16	3.33e-15	5 / 5	3.86e-04
SRP-dependent cotranslational protein targeting to membrane	21 / 113	0.01	1.11e-16	3.33e-15	5 / 5	3.86e-04
Cap-dependent Translation Initiation	22 / 120	0.011	1.11e-16	3.33e-15	13 / 18	0.001
Eukaryotic Translation Initiation	22 / 120	0.011	1.11e-16	3.33e-15	15 / 21	0.002
Keratinization	31 / 217	0.019	1.11e-16	3.33e-15	23 / 34	0.003
L13a-mediated translational silencing of Ceruloplasmin expression	22 / 112	0.01	1.11e-16	3.33e-15	2 / 3	2.31e-04
Eukaryotic Translation Termination	21 / 94	0.008	1.11e-16	3.33e-15	3 / 5	3.86e-04
Formation of the cornified envelope	31 / 129	0.011	1.11e-16	3.33e-15	16 / 27	0.002
Selenocysteine synthesis	21 / 94	0.008	1.11e-16	3.33e-15	2 / 7	5.40e-04
Response of EIF2AK4 (GCN2) to amino acid deficiency	21 / 102	0.009	1.11e-16	3.33e-15	4 / 16	0.001
Cellular responses to stress	40 / 565	0.049	1.11e-16	3.33e-15	46 / 227	0.018
Cellular responses to external stimuli	40 / 579	0.051	1.11e-16	3.33e-15	46 / 258	0.02
Developmental Biology	70 / 1,100	0.096	1.11e-16	3.33e-15	90 / 537	0.041
Influenza Infection	25 / 172	0.015	1.11e-16	3.33e-15	9 / 59	0.005
Selenoamino acid metabolism	21 / 118	0.01	1.11e-16	3.33e-15	2 / 33	0.003
Influenza Viral RNA Transcription and Replication	22 / 152	0.013	7.77e-16	2.18e-14	3 / 14	0.001

Supplementary 5. Tissue PVL samples top 5 pathways



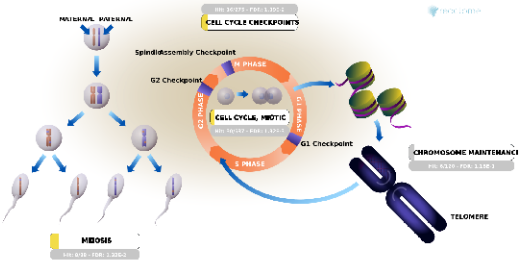
p-value = 9.03e-11

IMMUNE SYSTEM
106*



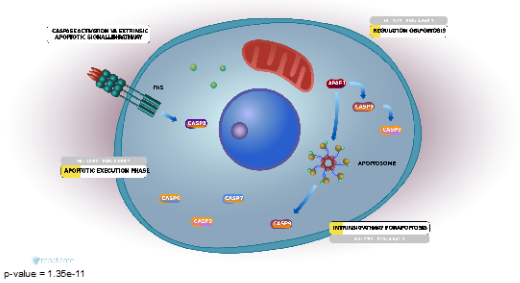
p-value = 1.21e-4

CELL CYCLE
35*



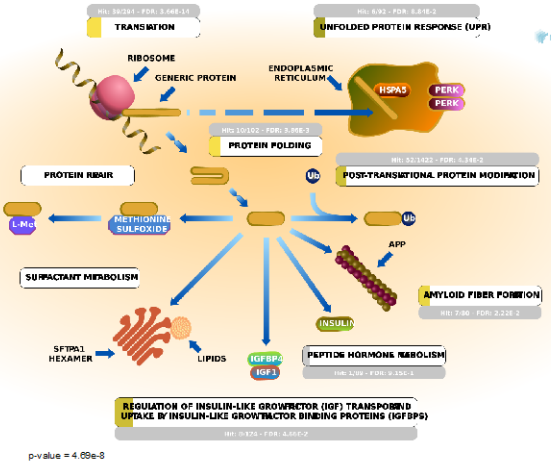
p-value = 6.71e-2

DNA REPLICATION
7*



p-value = 1.35e-11

APOPTOSIS
26*

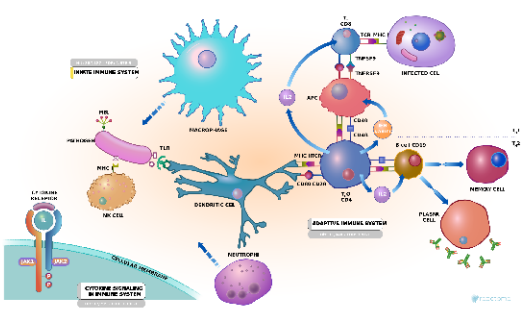


p-value = 4.69e-8

METABOLISM OF PROTEINS
95*

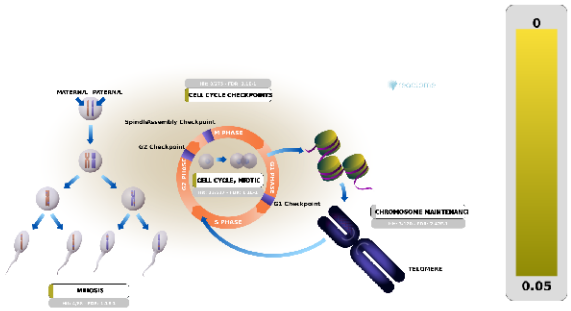
**Entities found in the analysis*
P value of the proteins represented in ISO diagrams of the reactome are indicated below them

Supplementary 6. Tissue OL samples top 5 pathways



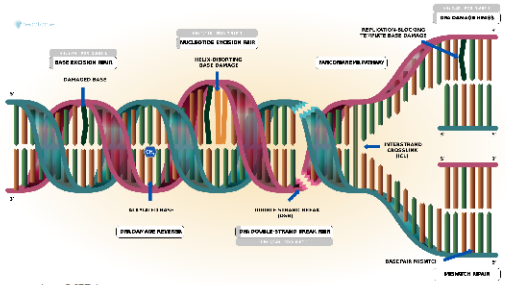
p-value = 2.7E-6

IMMUNE SYSTEM
56*



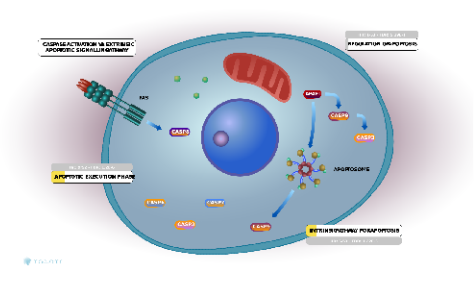
p-value = 1.46E-2

CELL CYCLE
16*



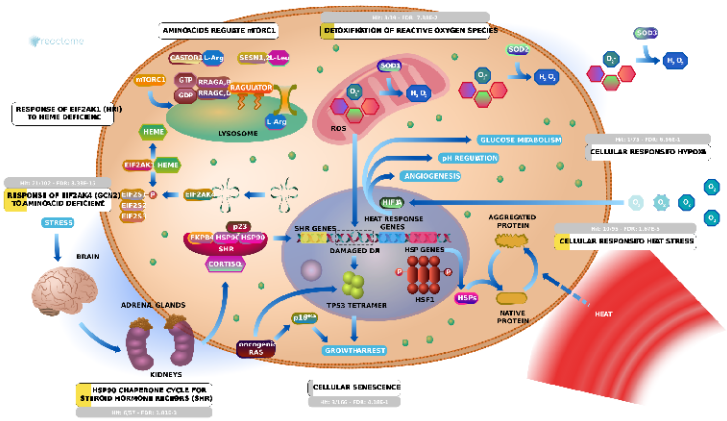
p-value = 2.87E-1

DNA REPAIR
6*



p-value = 8.09E-9

APOPTOSIS
16*

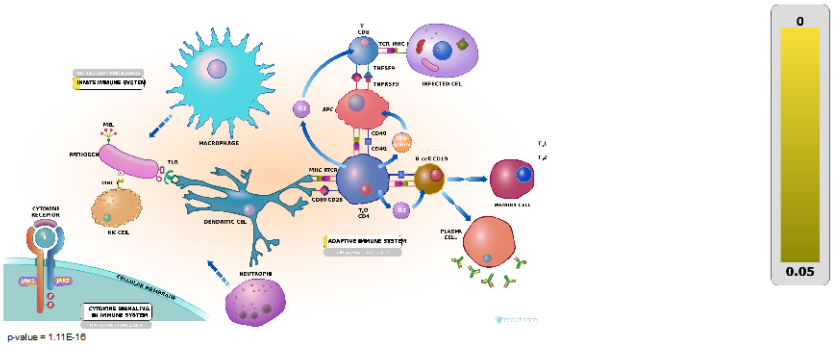


p-value = 1.11E-16

CELLULAR RESPONSES TO STRESS
40*

**Entities found in the analysis*
P value of the proteins represented in ISO diagrams of the reactome are indicated below them

Supplementary 7. Saliva PVL samples top 2 pathways

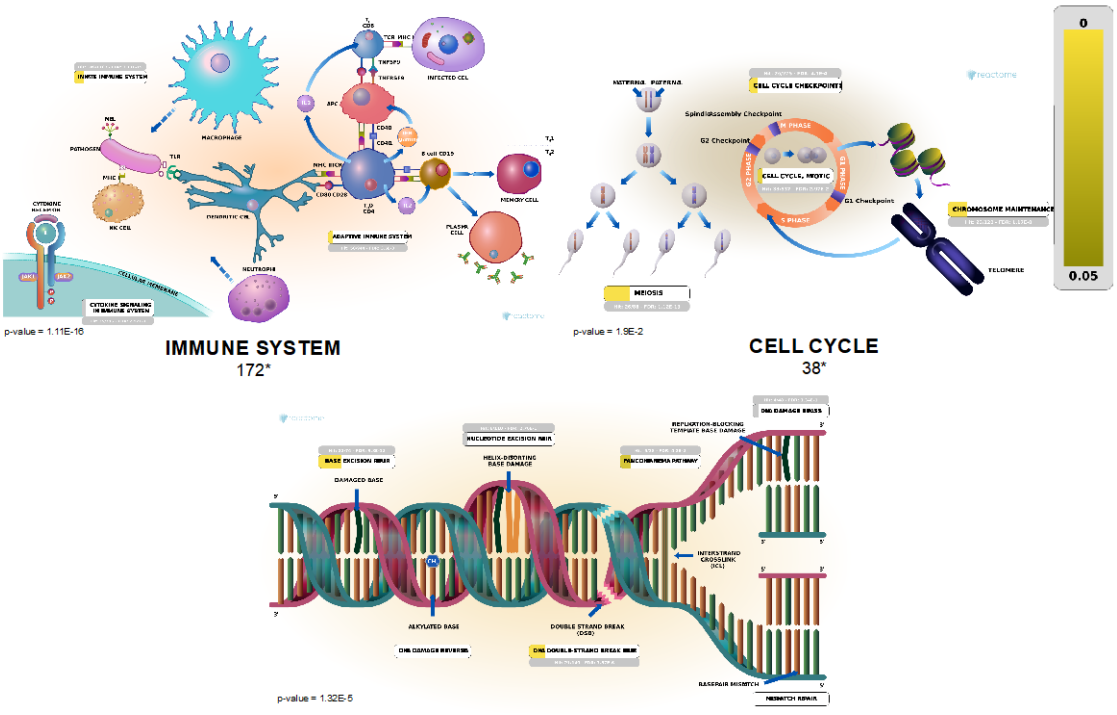


IMMUNE SYSTEM
159*

APOPTOSIS
12*

*Entities found in the analysis
P value of the proteins represented in ISO diagrams of the reactome are indicated below them

Supplementary 8. Saliva OL samples top 3 pathways



IMMUNE SYSTEM
172*

CELL CYCLE
38*

DNA REPAIR
30*

*Entities found in the analysis
P value of the proteins represented in ISO diagrams of the reactome are indicated below them

Supplementary 9. Pathway Analysis Report 3

The following table shows the 25 most relevant pathways sorted by p-value.

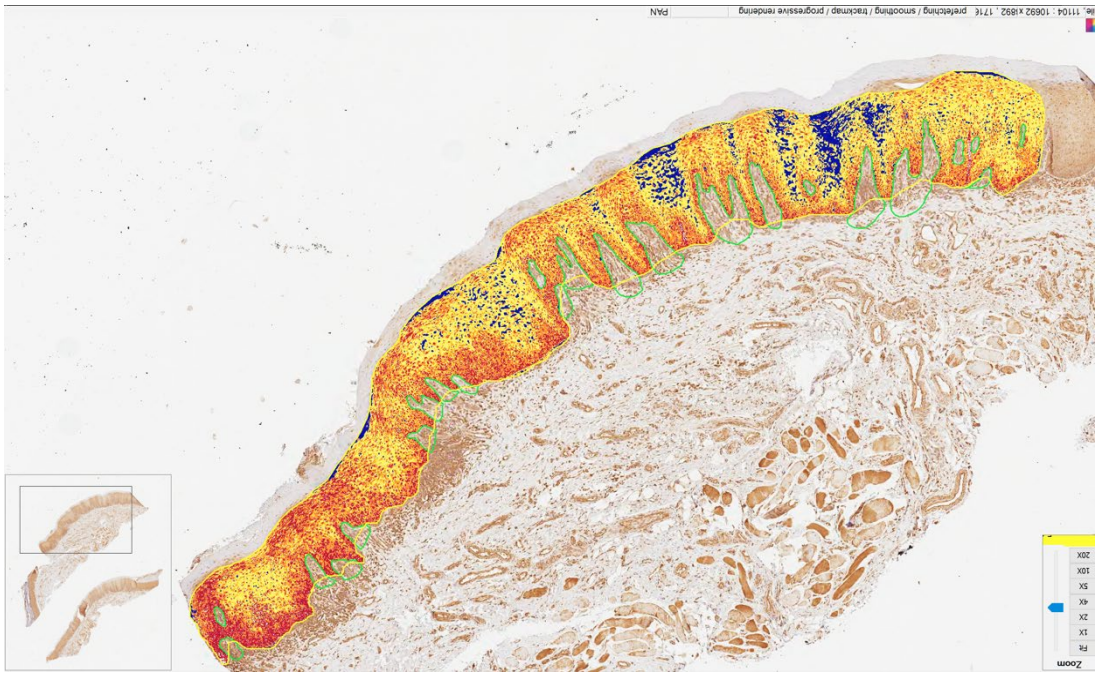
Pathway name	Entities				Reactions	
	found	ratio	p-value	FDR*	found	ratio
Scavenging of heme from plasma	34 / 99	0.009	1.11e-16	8.33e-15	12 / 12	9.26e-04
Classical antibody-mediated complement activation	27 / 95	0.008	1.11e-16	8.33e-15	2 / 2	1.54e-04
FCGR activation	28 / 101	0.009	1.11e-16	8.33e-15	6 / 6	4.63e-04
Neutrophil degranulation	80 / 480	0.042	1.11e-16	8.33e-15	10 / 10	7.71e-04
Regulation of Complement cascade	34 / 135	0.012	1.11e-16	8.33e-15	37 / 42	0.003
Complement cascade	34 / 146	0.013	1.11e-16	8.33e-15	59 / 71	0.005
CD22 mediated BCR regulation	25 / 70	0.006	1.11e-16	8.33e-15	3 / 4	3.09e-04
Initial triggering of complement	30 / 111	0.01	1.11e-16	8.33e-15	15 / 21	0.002
Platelet degranulation	33 / 128	0.011	1.11e-16	8.33e-15	7 / 11	8.49e-04
Regulation of actin dynamics for phagocytic cup formation	33 / 150	0.013	1.11e-16	8.33e-15	14 / 24	0.002
Binding and Uptake of Ligands by Scavenger Receptors	35 / 129	0.011	1.11e-16	8.33e-15	19 / 33	0.003
Response to elevated platelet cytosolic Ca2+	33 / 133	0.012	1.11e-16	8.33e-15	7 / 14	0.001
Innate Immune System	131 / 1,187	0.104	1.11e-16	8.33e-15	284 / 697	0.054
Immune System	165 / 2,398	0.21	1.11e-16	8.33e-15	470 / 1,634	0.126
Hemostasis	76 / 726	0.064	1.11e-16	8.33e-15	70 / 332	0.026
Creation of C4 and C2 activators	27 / 103	0.009	3.33e-16	2.33e-14	2 / 8	6.17e-04
Role of phospholipids in phagocytosis	28 / 114	0.01	4.44e-16	2.93e-14	5 / 12	9.26e-04
Leishmania phagocytosis	31 / 149	0.013	1.11e-15	6.00e-14	17 / 27	0.002
FCGR3A-mediated phagocytosis	31 / 149	0.013	1.11e-15	6.00e-14	17 / 27	0.002
Parasite infection	31 / 149	0.013	1.11e-15	6.00e-14	17 / 27	0.002
FCGR3A-mediated IL10 synthesis	29 / 128	0.011	1.11e-15	6.00e-14	10 / 20	0.002
Fcgamma receptor (FCGR) dependent phagocytosis	33 / 175	0.015	2.00e-15	1.02e-13	25 / 42	0.003
Antigen activates B Cell Receptor (BCR) leading to generation of second messengers	25 / 95	0.008	4.11e-15	2.01e-13	10 / 25	0.002
Role of LAT2/NTAL/LAB on calcium mobilization	25 / 102	0.009	1.98e-14	9.29e-13	6 / 7	5.40e-04
Cell surface interactions at the vascular wall	36 / 246	0.022	1.84e-13	8.26e-12	10 / 64	0.005

Supplementary 10. Pathway Analysis Report 4

The following table shows the 25 most relevant pathways sorted by p-value.

Pathway name	Entities				Reactions	
	found	ratio	p-value	FDR*	found	ratio
Scavenging of heme from plasma	27 / 99	0.009	1.11e-16	1.84e-14	10 / 12	9.26e-04
Neutrophil degranulation	70 / 480	0.042	1.11e-16	1.84e-14	8 / 10	7.71e-04
Platelet degranulation	29 / 128	0.011	1.11e-16	1.84e-14	6 / 11	8.49e-04
Response to elevated platelet cytosolic Ca2+	29 / 133	0.012	1.11e-16	1.84e-14	6 / 14	0.001
Innate Immune System	116 / 1,187	0.104	1.11e-16	1.84e-14	280 / 697	0.054
Immune System	141 / 2,398	0.21	1.11e-16	1.84e-14	431 / 1,634	0.126
Binding and Uptake of Ligands by Scavenger Receptors	27 / 129	0.011	2.22e-16	3.15e-14	15 / 33	0.003
Hemostasis	60 / 726	0.064	1.44e-15	1.79e-13	49 / 332	0.026
Regulation of Complement cascade	26 / 135	0.012	5.11e-15	5.67e-13	35 / 42	0.003
Complement cascade	26 / 146	0.013	3.03e-14	3.00e-12	57 / 71	0.005
Initial triggering of complement	23 / 111	0.01	4.52e-14	4.07e-12	15 / 21	0.002
Regulation of actin dynamics for phagocytic cup formation	26 / 150	0.013	5.58e-14	4.64e-12	12 / 24	0.002
CD22 mediated BCR regulation	19 / 70	0.006	7.47e-14	5.68e-12	3 / 4	3.09e-04
Leishmania phagocytosis	25 / 149	0.013	3.51e-13	2.18e-11	17 / 27	0.002
Parasite infection	25 / 149	0.013	3.51e-13	2.18e-11	17 / 27	0.002
FCGR3A-mediated phagocytosis	25 / 149	0.013	3.51e-13	2.18e-11	17 / 27	0.002
Platelet activation, signaling and aggregation	32 / 265	0.023	1.01e-12	5.85e-11	31 / 115	0.009
Classical antibody-mediated complement activation	20 / 95	0.008	1.60e-12	8.82e-11	2 / 2	1.54e-04
Fcgamma receptor (FCGR) dependent phagocytosis	26 / 175	0.015	1.73e-12	9.01e-11	23 / 42	0.003
Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs)	22 / 124	0.011	3.30e-12	1.62e-10	3 / 14	0.001
FCGR activation	20 / 101	0.009	4.75e-12	2.23e-10	6 / 6	4.63e-04
Creation of C4 and C2 activators	20 / 103	0.009	6.71e-12	3.02e-10	2 / 8	6.17e-04
Antigen activates B Cell Receptor (BCR) leading to generation of second messengers	19 / 95	0.008	1.39e-11	5.97e-10	10 / 25	0.002
Role of phospholipids in phagocytosis	20 / 114	0.01	3.96e-11	1.63e-09	5 / 12	9.26e-04

Supplementary 11. Example of Marking Intensity of Immunohistochemistry



4 CONCLUSÃO

A leucoplasia verrucosa proliferativa é a doença potencialmente maligna oral que mais acrimônia gera na clínica de estomatologia, pelo seu comportamento recalcitrante e pela alta taxa de transformação maligna que possui. Resulta evidente pensar que a leucoplasia verrucosa proliferativa é uma entidade clínica verdadeiramente diferente da leucoplasia oral convencional, não unicamente pelas características clinicas e histopatológicas, mas também pelos atributos moleculares que possui, os quais tem mostrado uma depleção marcada dos mecanismos de defesa mediados pela resposta imune inata e adaptativa, assim como a magnitude de expressão de proteínas ambivalentes que limitam e favorecem os processos de malignidade. Adicionalmente, constata-se que as diferenças moleculares expressas dentre as doenças potencialmente maligna orais, carcinoma espinho celular e controles saudáveis são o suficientemente nítidas para manifestar diferenças perceptíveis no proteoma salivar dos pacientes em questão.

REFERÊNCIAS*

1. Scully C, Bagan J. Oral squamous cell carcinoma overview. Oral Oncol. 2009; 45(4–5): 301–8.

2. El-Naggar AK, Chan JK, Grandis JR, Takata T, Slootweg PJ, editors. Oral potentially malignant disorders and oral epithelial dysplasia. In: WHO Classification of Head and Neck Tumours. 4th ed. Lyon: International Agency for Research on Cancer; 2017. p. 112–5.

3. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019; 144(8): 1941–53.

4. Sklenicka S, Gardiner S, Dierks EJ, Potter BE, Bell RB. Survival analysis and risk factors for recurrence in oral squamous cell carcinoma: does surgical salvage affect outcome? J Oral Maxillofac Surg. 2010; 68(6): 1270–5.

5. Van der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. Oral Oncol. 2009; 45(4–5): 317–23.

6. Van Der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; present concepts of management. Oral Oncol. 2010; 46(6): 423–5.

7. Neville BW, Damm DD, Allen CM, Chi AC. Oral and maxillofacial pathology. St. Louis, MO: Saunders/Elsevier; 2016. p. 355–63.

8. Iocca O, Sollecito TP, Alawi F, Weinstein GS, Newman JG, De Virgilio A, et al. Potentially malignant disorders of the oral cavity and oral dysplasia: a systematic review and meta-analysis of malignant transformation rate by subtype. Head Neck. 2020; 42(3): 539–55.

9. Petti S. Pooled estimate of world leukoplakia prevalence: a systematic review. Oral Oncol. 2003; 39(8): 770–80.

10. Cerero-Lapiedra R, Balade-Martinez D, Moreno-Lopez La, Esparza-Gomez G, Bagan Jv. Proliferative verrucous leukoplakia: a proposal for diagnostic criteria. Med Oral. 2010; e839–45.

11. Van der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. Oral Oncol. 2009; 45(4–5): 317–23.

* De acordo com o Guia de Trabalhos Acadêmicos da FOAr, adaptado das Normas Vancouver. Disponível no site da Biblioteca: http://www.foar.unesp.br/Home/Biblioteca/guia-de-normalizacao-atualizado.pdf

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

13. Villa A, Menon R, Kerr A, De Abreu Alves F, Guollo A, Ojeda D, et al. Proliferative leukoplakia: proposed new clinical diagnostic criteria. Oral Dis. 2018; 24(5): 749–60.

14. Celentano A, Glurich I, Borgnakke WS, Farah CS. World Workshop on Oral Medicine VII: Prognostic biomarkers in oral leukoplakia and proliferative verrucous leukoplakia: a systematic review of retrospective studies. Oral Dis. 2020; odi.13363.

15. Capella DL, Gonçalves JM, Abrantes AAA, Grando LJ, Daniel FI. Proliferative verrucous leukoplakia: diagnosis, management and current advances. Braz J Otorhinolaryngol. 2017; 83(5): 585–93.

16. Borgna SC, Clarke PT, Schache AG, Lowe D, Ho MW, McCarthy CE, et al. Management of proliferative verrucous leukoplakia: justification for a conservative approach. Head Neck. 2017; 39(10): 1997–2003.

17. Torrejon-Moya A, Jané-Salas E, López-López J. Clinical manifestations of oral proliferative verrucous leukoplakia: a systematic review. J Oral Pathol Med. 2020; 49(5): 404–8.

18. Hansen LS, Olson JA, Silverman S. Proliferative verrucous leukoplakia: a long-term study of thirty patients. Oral Surg Oral Med Oral Pathol. 1985; 60(3): 285–98.

19. Bagan JV, 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; 47(8): 732–5.

20. Gouvea Af, Moreira Ae, Reis Rr, Almeida Op, Lopes Ma. Proliferative verrucous leukoplakia, squamous cell carcinoma and axillary metastasis. Med Oral Patol Oral Cir Bucal. 2010; e704–8.

21. Scully C, Bagan J. Oral squamous cell carcinoma: overview of current understanding of aetiopathogenesis and clinical implications. Oral Dis. 2009; 15(6): 388–99.

22. Sklenicka S, Gardiner S, Dierks EJ, Potter BE, Bell RB. Survival analysis and risk factors for recurrence in oral squamous cell carcinoma: does surgical salvage affect outcome? J Oral Maxillofac Surg. 2010; 68(6): 1270–5.

23. van der Meij EH, Schepman K-P, van der Waal I. The possible premalignant character of oral lichen planus and oral lichenoid lesions: a prospective study. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2003; 96(2): 164–71.

24. González-Moles MÁ, Ruiz-Ávila I, González-Ruiz L, Ayén Á, Gil-Montoya JA, Ramos-García P. Malignant transformation risk of oral lichen planus: a systematic review and comprehensive meta-analysis. Oral Oncol. 2019; 96: 121–30.

25. Giuliani M, Troiano G, Cordaro M, Corsalini M, Gioco G, Lo Muzio L, et al. Rate of malignant transformation of oral lichen planus: a systematic review. Oral Dis. 2019; 25(3): 693–709.

26. Gillenwater AM, Vigneswaran N, Fatani H, Saintigny P, El-Naggar AK. Proliferative verrucous leukoplakia (PVL): a review of an elusive pathologic entity! Adv Anat Pathol. 2013; 20(6): 416–23.

27. Gillenwater AM, Vigneswaran N, Fatani H, Saintigny P, El-Naggar AK. Proliferative verrucous leukoplakia: recognition and differentiation from conventional leukoplakia and mimics. Head Neck. 2014; 36(11): 1662–8.

28. Lopes MA, Feio P, Santos-Silva AR, Vargas PA. Proliferative verrucous leukoplakia may initially mimic lichenoid reactions. World J Clin Cases. 2015; 3(10): 861–3.

29. Garcia-Pola M, Llorente-Pendas S, Gonzalez-Garcia M, Garcia-Martin J. The development of proliferative verrucous leukoplakia in oral lichen planus: a preliminary study. Med Oral. 2016; e328–34.

30. Müller S. Oral manifestations of dermatologic disease: a focus on lichenoid lesions. Head Neck Pathol. 2011; 5(1): 36–40.

31. Müller S. Oral lichenoid lesions: distinguishing the benign from the deadly. Mod Pathol. 2017; 30(s1): S54–67.

32. Fleskens SAJHM, Bergshoeff VE, Voogd AC, van Velthuysen M-LF, Bot FJ, Speel E-JM, et al. Interobserver variability of laryngeal mucosal premalignant lesions: a histopathological evaluation. Mod Pathol. 2011; 24(7): 892–8.

33. Kujan O, Oliver RJ, Khattab A, Roberts SA, Thakker N, Sloan P. Evaluation of a new binary system of grading oral epithelial dysplasia for prediction of malignant transformation. Oral Oncol. 2006; 42(10): 987–93.

34. Sperandio M, Brown AL, Lock C, Morgan PR, Coupland VH, Madden PB, et al. Predictive value of dysplasia grading and DNA ploidy in malignant transformation of oral potentially malignant disorders. Cancer Prev Res. 2013; 6(8): 822–31.

35. Upadhyaya JD, Fitzpatrick SG, Cohen DM, Bilodeau EA, Bhattacharyya I, Lewis JS, et al. Inter-observer Variability in the diagnosis of proliferative verrucous leukoplakia: clinical implications for oral and maxillofacial surgeon understanding: a collaborative pilot study. Head Neck Pathol. 2020; 14(1): 156–65.

36. Ghazali N, Bakri M, Zain R. Aggressive, multifocal oral verrucous leukoplakia: Proliferative verrucous leukoplakia or not? J Oral Pathol Med. 2003; 32: 383–92.

37. Carrard VC, Brouns EREA, van der Waal I. Proliferative verrucous leukoplakia; a critical appraisal of the diagnostic criteria. Med Oral Patol Oral Cir Bucal. 2013; 18(3): e411–3.

38. Bagan J, Scully C, Jimenez Y, Martorell M. Proliferative verrucous leukoplakia: a concise update. Oral Dis. 2010; 16(4): 328–32.

39. Dunn GP, Old LJ, Schreiber RD. The three es of cancer immunoediting. Annu Rev Immunol. 2004; 22(1): 329–60.

40. Dunn GP, Old LJ, Schreiber RD. The immunobiology of cancer immunosurveillance and immunoediting. Immunity. 2004; 21(2): 137–48.

41. Zitvogel L, Tesniere A, Kroemer G. Cancer despite immunosurveillance: immunoselection and immunosubversion. Nat Rev Immunol. 2006; 6(10): 715–27.

42. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. Nature. 2008; 454(7203): 436–44.

43. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. Science. 2011; 331(6024): 1565–70.

44. Farhi D, Dupin N. Pathophysiology, etiologic factors, and clinical management of oral lichen planus, part I: facts and controversies. Clin Dermatol. 2010; 28(1): 100–8.

45. Roopashree MR, Gondhalekar RV, Shashikanth MC, George J, Thippeswamy SH, Shukla A. Pathogenesis of oral lichen planus - a review: pathogenesis of oral lichen planus. J Oral Pathol Med. 2010; 39(10): 729–34.

46. Chitturi RT. Oral Lichen Planus: a review of etiopathogenesis, clinical, histological and treatment aspects. J Interdiscipl Med Dent Sci. 2014 [acesso 2020 Aug 6]; 02(05). Disponível em: https://www.omicsonline.com/open-access/oral-lichen-planus-a-review-of-etiopathogenesis-clinical-histological-and-jimds-1000142.php?aid=30035

47. Joseph BK. Oral cancer: prevention and detection. Med Principles Pract. 2002; 11: 32–35.

48. Farah CS, Jessri M, Bennett NC, Dalley AJ, Shearston KD, Fox SA. Exome sequencing of oral leukoplakia and oral squamous cell carcinoma implicates DNA damage repair gene defects in malignant transformation. Oral Oncol. 2019; 96: 42–50.

49. Farah CS, Fox SA. Dysplastic oral leukoplakia is molecularly distinct from leukoplakia without dysplasia. Oral Dis. 2019; 25(7): 1715–23.

50. Amado F, Lobo MJC, Domingues P, Duarte JA, Vitorino R. Salivary peptidomics. Expert Rev Proteomics. 2010; 7(5): 709–21.

51. Carnielli CM, Macedo CCS, De Rossi T, Granato DC, Rivera C, Domingues RR, et al. Combining discovery and targeted proteomics reveals a prognostic signature in oral cancer. Nat Commun. 2018; 9(1): 3598.

52. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The reactome pathway knowledgebase. Nucleic Acids Res. 2018; 46(D1): D649–55.

53. Indovina P, Marcelli E, Pentimalli F, Tanganelli P, Tarro G, Giordano A. Mass spectrometry-based proteomics: the road to lung cancer biomarker discovery. Mass Spectrom Rev. 2013; 32(2): 129–42.

54. Johann DJ, Rodriguez-Canales J, Mukherjee S, Prieto DA, Hanson JC, Emmert-Buck M, et al. Approaching solid tumor heterogeneity on a cellular basis by tissue proteomics using laser capture microdissection and biological mass spectrometry. J Proteome Res. 2009; 8(5): 2310–8.

55. Lorenzo-Pouso AI, Pérez-Sayáns M, Bravo SB, López-Jornet P, García-Vence M, Alonso-Sampedro M, et al. Protein-based salivary profiles as novel biomarkers for oral diseases. Dis Markers. 2018; 2018: 1–22.

56. Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. Nat Rev Clin Oncol. 2017; 14(9): 531–48.

57. Tyanova S, Cox J. Perseus: a bioinformatics platform for integrative analysis of proteomics data in cancer research. in: von Stechow L, editor. Cancer systems biology [Internet]. New York, NY: Springer New York; 2018 [acesso 2019 Nov 16]. p. 133–48. Disponível em: http://link.springer.com/10.1007/978-1-4939-7493-1_7

58. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods. 2016; 13(9): 731–40.

59. Kaufman E, Lamster IB. The diagnostic applications of saliva: a review. Crit Rev Oral Biol Med. 2002; 13(2): 197–212.

60. Gröschl M. Saliva: a reliable sample matrix in bioanalytics. Bioanalysis. 2017; 9(8): 655–68.

61. Aro K, Wei F, Wong DT, Tu M. Saliva liquid biopsy for point-of-care applications. front public health. 2017[acesso 2020 Abr 15]; 5. Disponível em: http://journal.frontiersin.org/article/10.3389/fpubh.2017.00077/full

62. Muetze T, Goenawan IH, Wiencko HL, Bernal-Llinares M, Bryan K, Lynn DJ. Contextual hub analysis tool (CHAT): a cytoscape app for identifying contextually relevant hubs in biological networks. [version 1; peer review: 1 approved]. F1000Research 2016, [acesso 2021 Jun 15] 5: 1745; 2016. Disponível em: (https://doi.org/10.12688/f1000research.9118.1)

63. Kuhn M, Szklarczyk D, Pletscher-Frankild S, Blicher TH, von Mering C, Jensen LJ, et al. STITCH 4: integration of protein–chemical interactions with user data. Nucleic Acids Res. 2014; 42(D1): D401–7.

64. Szklarczyk D, Santos A, von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. Nucleic Acids Res. 2016; 44(D1): D380–4.

65. Rintala M, Vahlberg T, Salo T, Rautava J. Proliferative verrucous leukoplakia and its tumor markers: systematic review and meta-analysis. Head Neck. 2019; 41(5): 1499–507.

Araraquara, 20 de janeiro de 2022.

Esteban Alexis Arroyo Ormeño