Biochemistry and Biophysics Reports 40 (2024) 101842

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

Label-based comparative proteomics of oral mucosal tissue to understand progression of precancerous lesions to oral squamous cell carcinoma

Vipra Sharma [a], Sundararajan Baskar Singh [a], Sabyasachi Bandyopadhyay [b], Kapil Sikka [c], Aanchal Kakkar [d], Gururao Hariprasad [a,*]

[a] Department of Biophysics, All India Institute of Medical Sciences, New Delhi, 110029, India
[b] Proteomics Sub-facility, Centralized Core Research Facility, All India Institute of Medical Sciences, New Delhi, 110029, India
[c] Department of Otorhinolaryngology, All India Institute of Medical Sciences, New Delhi, 110029, India
[d] Department of Pathology, All India Institute of Medical Sciences, New Delhi, 110029, India

ARTICLE INFO

Keywords:
Oral squamous cell carcinoma
Leukoplakia
Erythrolakia
Label-based proteomics
Dysplasia
Differential protein expression

ABSTRACT

Introduction: Oral squamous cell carcinomas typically arise from precancerous lesions such as leukoplakia and erythroplakia. These lesions exhibit a range of histological changes from hyperplasia to dysplasia and carcinoma in situ, during their transformation to malignancy. The molecular mechanisms driving this multistage transition remain incompletely understood. To bridge this knowledge gap, our current study utilizes label based comparative proteomics to compare protein expression profiles across different histopathological grades of leukoplakia, erythroplakia, and oral squamous cell carcinoma samples, aiming to elucidate the molecular changes underlying lesion evolution.

Methodology: An 8-plex iTRAQ proteomics of 4 biological replicates from 8 clinical phenotypes of leukoplakia and erythroplakia, with hyperplasia, mild dysplasia, moderate dysplasia; along with phenotypes of well differentiated squamous cell carcinoma and moderately differentiated squamous cell carcinoma was carried out using the Orbitrap Fusion Lumos mass spectrometer. Raw files were processed with Maxquant, and statistical analysis across groups was conducted using MetaboAnalyst. Statistical tools such as ANOVA, PLS-DA VIP scoring, and correlation analysis were employed to identify differentially expressed proteins that had a linear expression variation across phenotypes of hyperplasia to cancer. Validation was done using Bioinformatic tools such as ClueGO + Cluepedia plugin in Cytoscape to extract functional annotations from gene ontology and pathway databases.

Results and discussion: A total of 2685 protein groups and 12,397 unique peptides were identified, and 61 proteins consistently exhibited valid reporter ion corrected intensities across all samples. Of these, 6 proteins showed linear varying expression across the analysed sample phenotypes. Collagen type VI alpha 2 chain (COL6A2), Fibrinogen β chain (FGB), and Vimentin (VIM) were found to have increased linear expression across pre-cancer phenotypes of leukoplakia to cancer, while Annexin A7 (ANXA7) was seen to be having a linear decreasing expression. Collagen type VI alpha 2 chain (COL6A2) and Annexin A2 (ANXA2) had increased linear expression across precancer phenotypes of erythroplakia to cancer. The mass spectrometry proteomics data have been deposited to the ProteomeXchanger Consortium via the PRIDE partner repository with the data set identifier PXD054190. These differentially expressed proteins mediate cancer progression mainly through extracellular exosome; collagen-containing extracellular matrix, hemostasis, platelet aggregation, and cell adhesion molecule binding.

Conclusion: Label-based proteomics is an ideal platform to study oral cancer progression. The differentially expressed proteins provide insights into the molecular mechanisms underlying the progression of oral premalignant lesions to malignant phenotypes. The study has translational value for early detection, risk stratification, and potential therapeutic targeting of oral premalignant lesions and in its prevention to malignant forms.

* Corresponding author.
E-mail address: dr.hariprasadg@gmail.com (G. Hariprasad).

https://doi.org/10.1016/j.bbrep.2024.101842
Received 7 August 2024; Received in revised form 7 October 2024; Accepted 7 October 2024

1. Introduction

According to the most recent GLOBOCAN 2022 data from IARC, lip and oral cavity cancer holds the 16th position in global incidence and the 15th in mortality [1]. The disease is particularly prevalent in Asia, which has a 5-year prevalence rate of 62.4 %, and India stands out with a significantly high incidence [1,2]. Squamous cell carcinomas (SCCs), primarily arising from keratinocytes that constitute the majority of the oral cavity's epithelium, account for 90–95 % of head and neck cancers [3]. Oral squamous cell carcinomas (OSCC) primarily arise from precancerous oral potentially malignant disorders (OPMDs), yet they can occasionally surface in normal tissue, thereby increasing the risk of these lesions progressing into invasive oral cancers [4]. Leukoplakia and erythroplakia are the two primary types of premalignant lesions. While leukoplakia is, characterized by unscrapable white patches in the mouth, erythroplakia are fiery red soft velvety oral lesions [5,6]. Prevalence of oral leukoplakia is 1.39 %, with a worldwide estimate of 3.41 %, and higher rates among males, those over 60, smokers, and alcohol consumers [7]. On the other hand, prevalence of oral erythroplakia is comparatively lower, estimated to be 0.17 % from a minimal 0.13 % to a substantial 40.8 % [8]. Notably, oral erythroplakia presents a significantly higher propensity for malignant transformation, with a rate of 19.9 % [9].

The development of an invasive cancer is thought to arise from stepwise progression through increasing degrees of mucosal dysplasia. The histological transition of oral epithelial cells is a manifestation of the underlying molecular and cellular changes that occur during the development of oral cancer. These changes are grouped as genomic, proliferation and differentiation categories [10]. Overexpression of EGFR, correlated with poor prognosis, is observed in malignant, premalignant, and normal-appearing tissues from squamous cell carcinoma of head and neck (SCCHN) patients, while TGF-α staining in the stratum germinativum and EGFR staining in the stratum spinosum both increase linearly with the severity of dysplasia in oral leukoplakia compared to control mucosa [10–12]. The immunohistochemical analysis of various oral tissue specimens revealed (i) Loss of p16 is an early event in oral premalignant lesions and SCCHN tumors; (ii) Rb loss is associated with the transition from hyperplasia to dysplasia; (iii) Rb loss with cyclin D1 or p53 overexpression is linked to the progression from dysplasia to SCCHN, and (iv) Rb loss and p53 overexpression predict poor survival in SCCHN [13]. Though these single protein quantitation sheds light on cancer progression, an unbiazed quantitative proteomic approach should help in better understanding of the molecular mechanisms that operate in dysplasia to neoplasia transformation.

In the realm of oral oncology, high-throughput quantitative proteomics has been instrumental in identifying biomarkers within oral epithelium, from morphologically malignant cells to epithelial dysplasia and adjacent normal cells, providing valuable insights for the early detection and management of oral cancer [14]. Protein expression discrepancies between dysplastic leukoplakia and histological normal oral tissues, as well as among variously differentiated OSCC types too have been addressed [15,16]. However, there remains a gap in our understanding of the molecular cascade driving the transition from normal epithelium through various grades of dysplasia to OSCC. In the recent past our lab has been looking into protein profiling to explain prediction of precancer lesions to cancer, and also to explain the role of tobacco in oral cancer [17,18].

Our current study aims to discern protein biomarkers that mark the multistage progression to squamous cell carcinoma using 8-plex iTRAQ proteomics experiment. With reporter ion-based quantification, we aim to reduce sample variability and leverage multiplexing for precise proteomic profiling. This study compares hyperplastic leukoplakia and erythroplakia cases against those with progressive dysplasia and OSCC, delineating the molecular milestones of lesion evolution.

2. Methodology

2.1. Ethics, patient screening, and sample collection

The study received ethical approval from the Institute Ethics Committee at the All India Institute of Medical Sciences, New Delhi, India (Ref. No. IECPG-370/26.08.2020). A brief overview of the methodology is provided in Fig. 1. All procedures were executed in compliance with the guidelines of the Helsinki Declaration. Patients with clinical manifestations of leukoplakia, erythroplakia, and oral squamous cell carcinoma were screened and diagnosed at the Out-Patient Department of Otorhinolaryngology. Prior to participation, informed consent was obtained from each patient. Biopsies were performed on the suspected lesions. The specimens were then divided, with one part sent to the Department of Pathology for histopathological evaluation and the other part cryopreserved at −80 °C for future proteomic analysis at the Proteomics Facility.

2.2. Patient inclusion and exclusion criteria

The study included patients with oral cavity lesions. Exclusion criteria were those who had previously undergone neoadjuvant chemotherapy or radiotherapy, and patients with concurrent morbidities or infections.

2.3. Histopathological evaluation

Hematoxylin and eosin-stained slides were reviewed to verify morphological diagnoses, adhering to the World Health Organization's 2022 criteria for oral epithelial dysplasia and Broder's grading for squamous cell carcinoma. Collected samples were then annotated based on oral clinical features combined with histopathological findings.

2.4. Protein extraction

Tissue specimens from oral cavity lesions were finely minced in lysis buffer (8 M urea, 2 M thiourea, 4 % CHAPS), followed by vortexing. Ultrasonication was performed at 45 Hz and 90 W for 20 min. Post-ultrasonication, the lysate was centrifuged at 12,000 rpm for 15 min at 4 °C, and the supernatant was retrieved. A 2D clean-up kit was employed to purify the protein extract as per the manufacturer's protocol. The pellet obtained was dissolved in 8 M urea/500 mM TEAB (pH 8), and protein concentration was quantified at 280 nm.

2.5. In-Solution tryptic digestion and desalting

Each proteome sample (25 μg) was reduced with 10 mM dithiothreitol (DTT) at 37 °C for 30 min, followed by alkylation with 50 mM iodoacetamide (IAA) in the dark at room temperature for 20 min. The lysate was then diluted with 500 mM triethyl ammonium bicarbonate (TEAB) (pH 8.0) to decrease the urea concentration below 1 M. Trypsin was added at a 1:50 (w:w) ratio, and the mixture was digested overnight at 37 °C. The digested samples were lyophilized and reconstituted in 200 μL of 0.1 % trifloroacetic acid (TFA) and 2 % acetonitrile (ACN), followed by vigorous vortexing. Desalting was conducted using an Oasis HLB 96-well μElution plate (30 μm, 5 mg, Waters, Milford, MA) prewashed with ACN and equilibrated with 0.1 % TFA. After discarding the flow-through and washing with 0.1 % TFA and HPLC-grade water, peptides were eluted with 70 % ACN and vacuum dried.

2.6. iTRAQ labelling and high pH reversed-phase peptide fractionation

Digested peptides were labeled with iTRAQ 8-plex reagents, each correlating to distinct phenotypes: 113 (leukoplakia with hyperplasia), 114 (erythroplakia with hyperplasia), 115 (leukoplakia with mild dysplasia), 116 (erythroplakia with mild dysplasia), 117 (leukoplakia

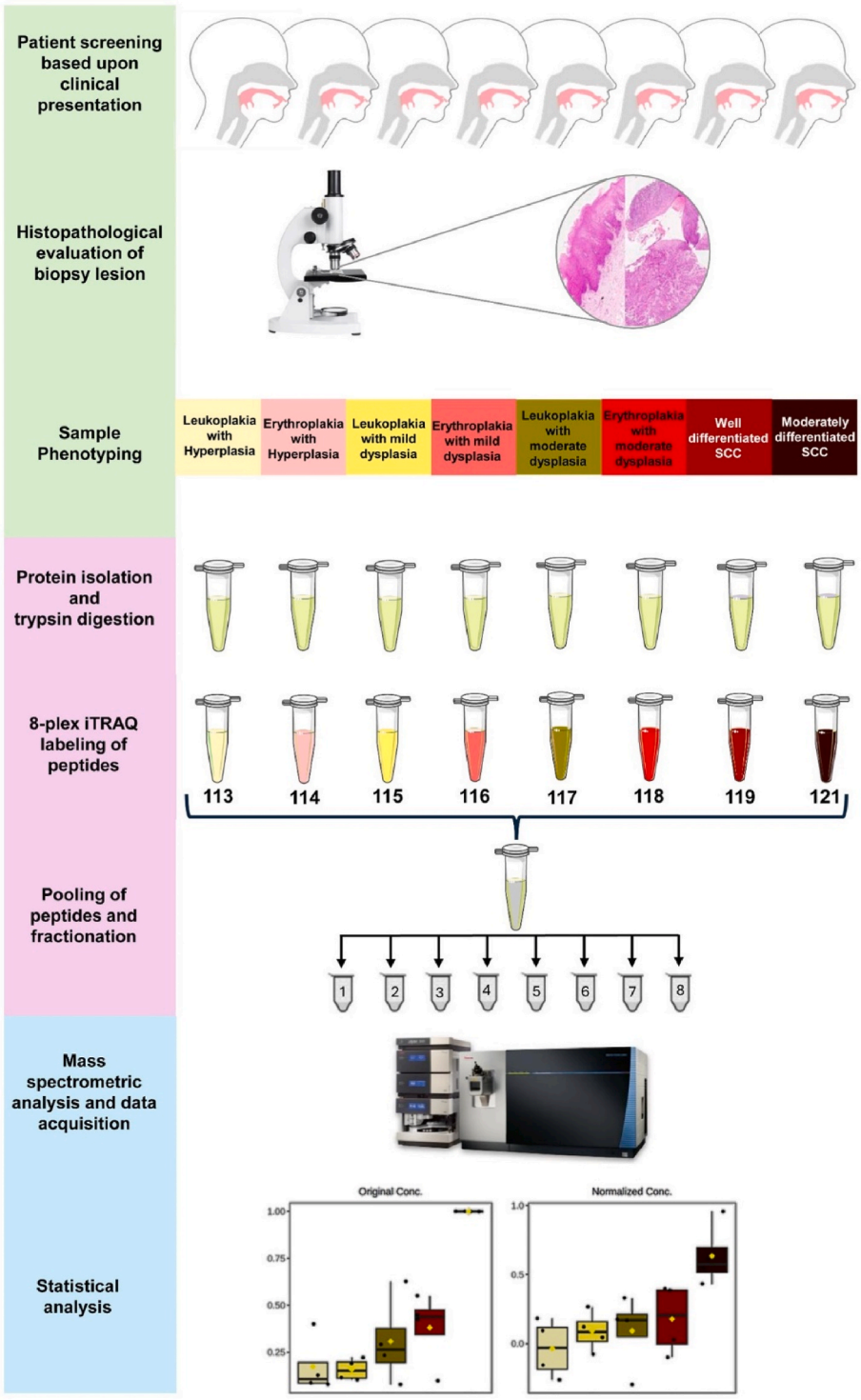

Fig. 1. Workflow of the methodology used in the iTRAQ 8-plex quantitative proteomic experiment.

with moderate dysplasia), 118 (erythroplakia with moderate dysplasia), 119 (well differentiated SCC), and 121 (moderately differentiated SCC). The labeling was performed following the manufacturer's protocol (AB Sciex, Foster City, CA, USA) by reconstituting each iTRAQ tag in ethanol and adding it to the respective digested sample in 0.5 M TEAB. After a 2-h room temperature incubation, the labeled samples were combined and dried. High pH reversed-phase fractionation using spin column (Thermo Fisher Scientific, USA) was then applied to the pooled iTRAQ-labeled peptides, separating them into eight fractions of increasing ACN concentrations. Fractions were pooled from three spin columns per biological replicate to maximize peptide recovery, vacuum dried, and stored for subsequent mass spectrometry analysis.

2.7. Mass spectrometric data acquisition

All spectra were acquired on an Oribtrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific,USA) coupled to an Easy-nLC 1200 (Thermo Fisher Scientific,USA) ultra-high pressure liquid chromatography (UHPLC) pump. Peptides were loaded onto Acclaim PepMap 100 C18 trap column (3 μm particle size, 20 mm length x 75 μm I.D.; Thermo

Fisher Scientific,USA) and then separated on an Easy-Spray PepMap RSLC C18 column (2 μm particle size, 150 mm length × 75 μm I.D.; Thermo Fisher Scientific,USA) with a gradient consisting of 0%–45 % (80 % ACN, 0.1 % FA) at 300 nL/min and injected for MS analysis The scan sequence began with FTMS1 spectra (RF lens = 30 %; Resolution = 120,000 (at *m/z* 200); Scan range = 350–1800 *m/z*; Automatic gain control (AGC) target = 100 %, Max injection time = 50 ms). MS2 spectra were collected in data dependent mode for top ten intense peptides, that were sequentially isolated in the quadrupole with an isolation window of 0.7 and fragmented using higher-energy collisional dissociation (HCD) with a normalized collision energy of 38 %. The MS2 scans were performed at a cycle time of 3 s. The MS2 fragment ions were detected in the Orbitrap mass analyzer at a resolution of 50,000 (at *m/z* 200). The normalized automatic gain control (AGC) target was set to 250 % to maximize the number of ions accumulated. The MS2 spectra were recorded in centroid mode.

2.8. Quantification and downstream analyses

Raw MS files were processed using MaxQuant software (2.5.1.0). The Andromeda search engine was employed to search against the *Homo sapiens* Uniprot database (downloaded on May 31, 2024) and common contaminants (provided with MaxQuant). Since the dataset lacks reference channels, all channels were treated as reference channels, and their total signal sum was used. Reporter ion MS2 was selected as the "type" under "group-specific parameters." We used 8plex iTRAQ as the isobaric labeling method providing correction factor for the isotopic impurities of the different iTRAQ reagents according to manufacturer specifications. Isobaric matching between runs was applied using the "MBR" option in the "identification" section under "global parameters", and normalization was performed as a weighted ratio to the reference channel [19]. The search was configured with fixed modification of carbamidomethyl cysteine and dynamic modifications of methionine oxidation and N-terminal acetylation. The FDR was set to <0.01 for peptide and protein identification. All nanoLC-MS/MS raw files and MaxQuant search results were deposited at the ProteomeXchange Consortium via the PRIDE partner repository data set identifier PXD054190 [20,21]. The file "Proteingroups.txt" contained corrected reporter ion intensities for the identified proteins. To perform further statistical analyses and downstream processing, MetaboAnalyst (v6.0) was utilized [22]. To ensure reliable comparisons and accurate analyses, the original protein expression intensities underwent filtering. Additionally, the expression values were normalized using the median value of each sample and transformed by log transformation with base 10.

This subset of proteins was subjected to further statistical analysis such as (i) univariate analysis using one-way ANOVA, (ii) multivariate analysis employing Partial Least Squares Discriminant Analysis (PLS-DA) and, (iii) correlation analysis to identify the patterns of protein expression changes that occur linearly during the disease progression.

Univariate statistical analysis was performed by one-way ANOVA followed by Tukey's multiple comparison test at a cut-off p-value <0.05. Multivariate Partial Least Square Discriminant Analysis (OPLS-DA) was applied to discriminate the different histological grades. Proteins having VIP scores greater than 1.5 were considered significant for the final model. Pattern hunter feature was used to perform correlation analysis against a predefined pattern to identify protein expression changes during disease progression - from hyperplasia to dysplasia to neoplasia - ranked by Pearson correlation. ClueGO (v2.5.10) and CluePedia (v1.5.10) plugins on Cytoscape version 3.10.0 were used to perform functional enrichment analysis and visualize the non-redundant biological terms for large clusters of proteins identified in four biological replicates of hyperplasia and moderately differentiated squamous cell carcinoma samples (MDSCC) [23]. A two-sided hypergeometric test with a p-value cut-off of <0.05 was employed, with the kappa score threshold set to 0.4 and a GO term/pathway network created using a GO tree interval of 3–8 levels. Only pathways with a p-value <0.05 were displayed in the functionally grouped network. Differentially expressed proteins and their direct first neighbors were extracted from the complex network to gain protein-specific insights under the hyperplasia and MDSCC conditions.

The resulting set of proteins obtained from one-way ANOVA, PLS-DA VIP scores, and correlation analysis were subjected to a Venn diagram analysis. This integrated statistical approach utilized the following criteria: One-way ANOVA with p-value <0.05 to identify proteins with significant mean differences between two or more groups. PLS-DA VIP scores >1.5 to select the most important proteins for discriminating the sample groups. Pearson correlation analysis to identify proteins based upon correlation coefficient (r): $0.6 \leq r \leq 1$ (positive correlation) and $-0.6 \leq r \leq 1$ (negative correlation).

3. Results and Discussion

3.1. Patient recruitment

In the discovery phase of our study, we recruited thirty-two patients, each with four biological replicates per phenotype. The clinical profile of these patients are documented in Table 1. All study participants were male, with the majority having a history of tobacco consumption (either through chewing or smoking). Additionally, some patients reported alcohol consumption. Notably, this prevalence of male-dominated cohort and tobacco being major etiological factor for both precancerous and cancerous oral lesions align with previous studies [24,25]. Due to ethical constraints, normal/healthy oral cavity tissue samples could not be obtained. Instead, we utilized epithelial hyperplastic tissues from leukoplakia and erythroplakia as controls. Although dysplasia is typically categorized into three grades, we omitted the severe dysplasia based on repeat biopsies that consistently revealed differentiated carcinoma in the majority of cases. We also encountered challenges in obtaining sufficient number of samples of poorly differentiated squamous cell carcinoma due to its rarity. This limitation aligns with findings from previous research [26].

3.2. Histopathological phenotyping

Histopathological analysis has been used as the 'Gold Standard' for the phenotyping of the oral mucosal tissue samples that were collected. Representative histopathological profiles of biopsy tissues from leukoplakia, erythroplakia, and cancer are provided in Fig. 2. These architectural and cytological changes in the oral epithelium across different phenotypes are: Hyperplasia: There is an increase in the thickness of the epithelium without loss of polarity or irregular stratification of the epithelial cells. Nuclear hyperchromasia and pleomorphism are absent (Fig. 2A and F); Mild dysplasia: There is an increase in epithelial thickness accompanied by disorganization of basal cells, which show mild nuclear hyperchromasia and pleomorphism. These changes are restricted to the lower one-third of the epithelium (Fig. 2B and G); Moderate dysplasia: There is an expanded proliferative compartment with drop shaped rete-ridges. The epithelial cells have an increased nucleus: cytoplasm ratio, enlarged hyperchromatic nuclei with considerable variation in size and shape. Mitotic activity is notably increased (Fig. 2C and H); Well differentiated squamous cell carcinoma: There is an increased cytoplasmic keratinization and formation of keratin pearls (Fig. 2D and I); Moderately differentiated squamous cell carcinoma: There is a greater cellular and nuclear pleomorphism. Atypical mitoses are present (Fig. 2E and J).

3.3. Protein profiling

Following mass spectrometric analysis of 32 oral cavity lesion samples, a total of 2685 protein groups and 12,397 unique peptides were identified. To ensure robust and reliable analysis, proteins identified by site, reverse decoys, and potential contaminants were excluded from

Table 1
Clinical profile of patients recruited for the study.

Sr. No.	Patient ID	Age/ Gender	Habits	Clinical Diagnosis	Histopathology	Phenotypes	iTRAQ label
1	DKM23	53 y/M	Tobacco chewing	Leukoplakic lesion over right buccal mucosa	Hyperplasia	Leukoplakia with hyperplasia	113
2	HC23	38 y/M	Tobacco chewing	Indurated leukoplakic lesion over right buccal mucosa	Hyperplasia	Leukoplakia with hyperplasia	113
3	ZA23	39 y/M	Tobacco chewing	Leukoplakic lesion in left buccal mucosa	Hyperplasia	Leukoplakia with hyperplasia	113
4	MJ22	42 y/M	Tobacco chewing + Smoking	Leukoplakic patch on left side of hard palate	Hyperplasia	Leukoplakia with hyperplasia	113
5	PKB23	40 y/M	Tobacco chewing	Erythroplakic patch + OSMF on left buccal mucosa	Hyperplasia	Erythroplakia with hyperplasia	114
6	VKS23	44 y/M	Tobacco chewing	Erythroplakic lesion in left buccal mucosa	Hyperplasia	Erythroplakia with hyperplasia	114
7	HPY23	39 y/M	Tobacco chewing	Erythroplakic lesion in left buccal mucosa	Hyperplasia	Erythroplakia with hyperplasia	114
8	DS22	53 y/M	Tobacco chewing + Smoking	Erythroplakic lesion in left buccal mucosa	Hyperplasia	Erythroplakia with hyperplasia	114
9	RS19	34 y/M	Smoking	Leukoplakia on left retro molar trigone	Mild dysplasia	Leukoplakia with mild dysplasia	115
10	V20	45 y/M	Tobacco chewing	Leukoplakia on left buccal mucosa	Mild dysplasia	Leukoplakia with mild dysplasia	115
11	AN23	28 y/M	Smoking	Leukoplakia on left buccal mucosa	Mild dysplasia	Leukoplakia with mild dysplasia	115
12	OP23	34 y/M	Tobacco chewing	Leukoplakic lesion on left buccal mucosa	Mild dysplasia	Leukoplakia with mild dysplasia	115
13	PKS22	33 y/M	Tobacco chewing	Erythroplakic patch on right buccal mucosa	Mild dysplasia	Erythroplakia with mild dysplasia	116
14	DKS23	42 y/M	Tobacco chewing	Erythroplakic patch on right buccal mucosa	Mild dysplasia	Erythroplakia with mild dysplasia	116
15	D23	59 y/M	Tobacco chewing	Erythroplakic patch on right gingivobuccal mucosa	Mild dysplasia	Erythroplakia with mild dysplasia	116
16	D24	40 y/M	Tobacco chewing	Erythroplakic patch on left buccal mucosa	Mild dysplasia	Erythroplakia with mild dysplasia	116
17	BB20	71 y/M	Information not available	Indurated leukoplakia on right buccal mucosa	Moderate dysplasia	Leukoplakia with moderate dysplasia	117
18	R20	40 y/M	Information not available	Leukoplakia on right buccal mucosa	Moderate dysplasia	Leukoplakia with moderate dysplasia	117
19	JKS23	60 y/M	Tobacco chewing	Leukoplakic lesion in left buccal mucosa	Moderate dysplasia	Leukoplakia with moderate dysplasia	117
20	HS20	54 y/M	Tobacco chewing + Smoking	Leukoplakia on right buccal mucosa	Moderate dysplasia	Leukoplakia with moderate dysplasia	117
21	VS21	53 y/M	No addiction	Erythroplakia in retromolar trigone	Moderate dysplasia	Erythroplakia with moderate dysplasia	118
22	MD22	50 y/M	Information not available	Erythroplakia on buccal mucosa	Moderate dysplasia	Erythroplakia with moderate dysplasia	118
23	RK23	45 y/M	Tobacco chewing	Erythroplakia with intermittent white patch in right mucosa	Moderate dysplasia	Erythroplakia with moderate dysplasia	118
24	PKG23	53 y/M	Tobacco chewing + smoking + alcohol consumption	Erythrematous lesion at from left buccal mucosa	Moderate dysplasia	Erythroplakia with moderate dysplasia	118
25	S19	55 y/M	Information not available	Indurated ulceric lesion over left buccal mucosa	Well differentiated SCC	Well differentiated SCC	119
26	P22	63 y/M	Smoking	Ulceroproliferative growth over upper alveolus	Well differentiated SCC	Well differentiated SCC	119
27	SKP20	56 y/M	Smoking	Ulcerative lesion over anterior left lateral border of tongue	Well differentiated SCC	Well differentiated SCC	119
28	I20	35 y/M	Tobacco chewing	Ulceroproliferative growth on right buccal mucosa	Well differentiated SCC	Well differentiated SCC	119
29	HOS21	38 y/M	Tobacco chewing	Ulceroproliferative growth over lower gingivobuccal sulci	Moderately differentiated SCC	Moderately differentiated SCC	121
30	MLR22	53 y/M	Tobacco chewing + alcohol consumption	Non- healing ulcer in right buccal mucosa	Moderately differentiated SCC	Moderately differentiated SCC	121
31	O21	48 y/M	Smoking	Ulceroproliferative growth on left retromolar trigone	Moderately differentiated SCC	Moderately differentiated SCC	121
32	RCB20	56 y/M	Smoking	Ulceroproliferative growth on left buccal mucosa	Moderately differentiated SCC	Moderately differentiated SCC	121

further consideration. The investigation was focused on a subset of 61 proteins that consistently exhibited valid reporter ion corrected-intensity across all histological grades of samples (leukoplakia; 113, 115, 117, 119, and 121; and erythroplakia: 114, 116, 118, 119, and 121). This subset of proteins was subjected to further statistical analysis that is discussed below.

Univariate statistical analysis performed using one-way ANOVA followed by Tukey's multiple comparison test resulted in proteins that changed significantly among different histological grades of leukoplakia and erythroplakia. The analysis revealed notable differences, with a p-value threshold of <0.05. 9 proteins that showed statistical significance across leukoplakia histopathology are Eukaryotic elongation factor 2 kinase (EEF2), Collagen type VI alpha 2 chain (COL6A2), Heat shock protein beta-1 (HSPB1), Annexin A2 (ANXA2), Collagen type VI alpha 1

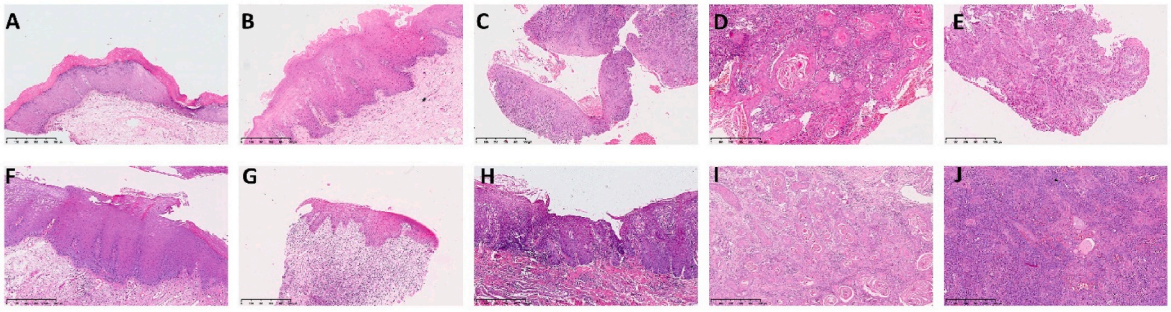

Fig. 2. Hematoxylin and eosin-stained histopathology sections of biopsies. Upper panel shows biopsies from leukoplakic lesions: (A) epithelial hyperplasia, (B) mild dysplasia, (C) moderate dysplasia, (D) well-differentiated squamous cell carcinoma, and (E) moderately differentiated squamous cell carcinoma. Lower panel shows bopsies from erythroplakic lesions: (F) epithelial hyperplasia, (G) mild dysplasia, (H)moderate dysplasia, (I) well differentiated squamous cell carcinoma, and (J) moderately differentiated squamous cell carcinoma.

chain (COL6A1), Annexin A2 (ANXA7), Prolargin (PRELP), Fibrinogen β chain (FGB) and Vimentin (VIM), while 7 proteins that exhibited significance across erythroplakia histopathology are Eukaryotic elongation factor 2 kinase (EEF2), Collagen type VI alpha 2 chain (COL6A2), Fatty acid-binding protein 5 (FABP5), CNDP2, Annexin A2 (ANXA2), Heat shock protein beta-1 (HSPB1) and Cytochrome *c* oxidase subunit 5A, mitochondrial (COX5A).

Multivariate Partial Least Square Discriminant Analysis (PLS-DA) was applied to discriminate the proteins responsible for separation of different histopathological grades of leukoplakia and erythroplakia, two common premalignant oral lesions. As illustrated in Fig. 3A and B, the principal components extracted by partial least squares-discriminant analysis (PLS-DA) explained co-variance of 35.2 % and 34.5 % for leukoplakia and erythroplakia respectively. The classification did not reveal distinct clustering of the different histopathological grades in the case of leukoplakia. Instead, we observed a substantial overlap between the initial grades of histopathology and the successive, more advanced grades. A similar pattern was observed for the erythroplakia samples, with the exception that the moderately differentiated carcinoma group formed a distinct cluster, while hyperplasia showed overlap with mild dysplasia, and moderate dysplasia overlapped with well-differentiated carcinoma. These findings suggest that, despite the distinct histopathological grading, there are underlying molecular events and heterogeneity within the non-tumor and tumor microenvironments that govern the differences in protein expression and interactions. The lack of clear separation between the histopathological grades, particularly in leukoplakia, highlights the complex and dynamic nature of the molecular landscape in premalignant oral lesions, where factors beyond the

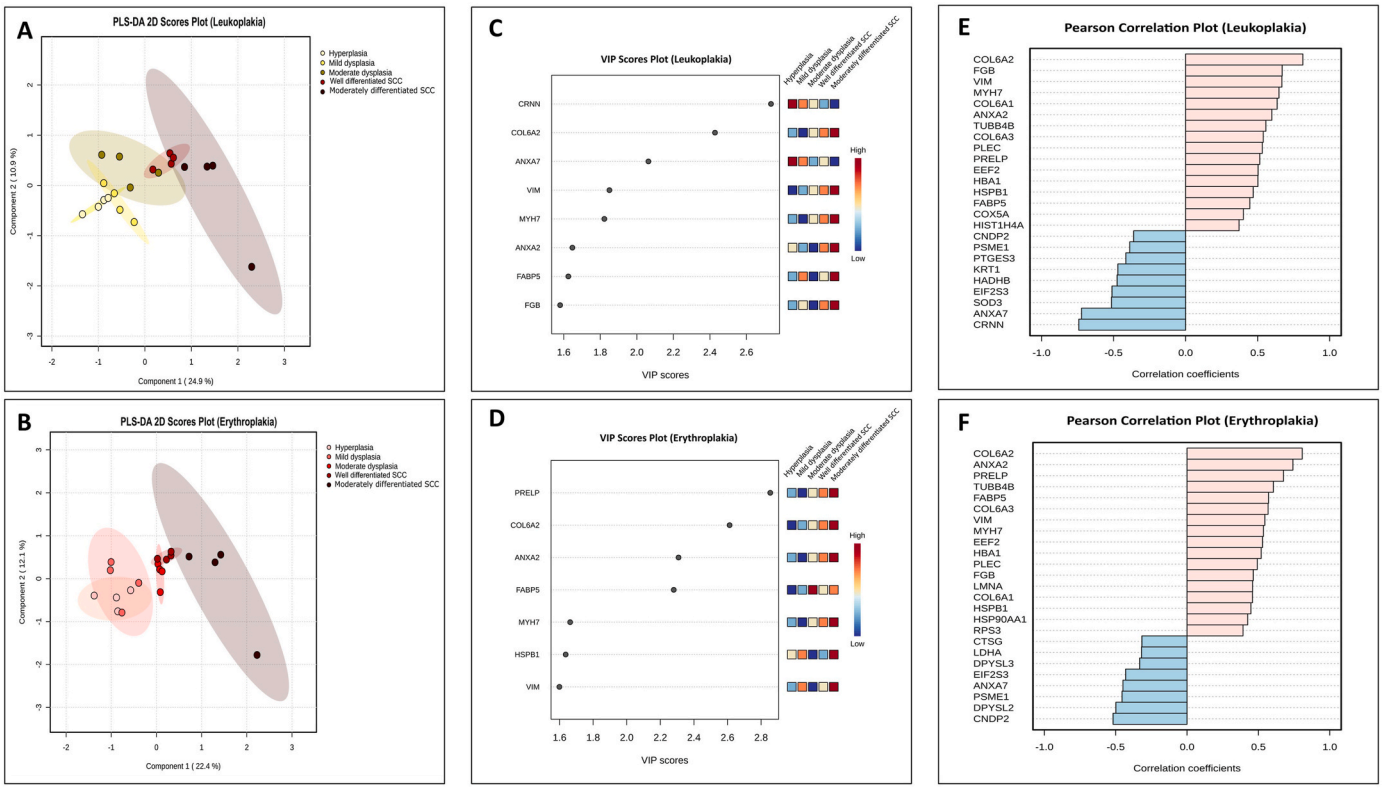

Fig. 3. Statistical analysis on all proteins identified in leukoplakia and erythroplakia. (A,B) PLS-DA 2D scores plots showing the extent of separation among different histological grades; (C,D) PLS-DA VIP score plots mapping the proteins responsible for group separation, ranked by their importance (VIP score) based upon component 1 in discriminating the sample groups. The colored boxes represent the relative protein intensity in each group, with red indicating high and blue indicating low intensity. (E,F) Pearson correlation analysis of the top 25 proteins, mapping the proteins that are positively correlated (light pink) and negatively correlated (light blue) across the worsening histopathological grades of the lesions.

histological appearance may play a crucial role in driving the transition from precancerous to cancerous states.

A key component of the PLS-DA analysis is the Variable Importance in Projection (VIP) score which estimates the degree of contribution of each variable (protein) in the model's ability to classify or discriminate between the predefined groups. Fig. 3C and D illustrates VIP scored proteins and relative expression levels of corresponding protein in progressive grades. CRNN and PRELP exhibit the highest VIP scores in leukoplakia and erythroplakia, respectively. Therefore, contributing significantly to separation between histopathological grades. Our findings align with a previous study, demonstrating that Cornulin (CRNN) expression decreases progressively as dysplasia and OSCC grades increase [27].

Correlation analysis using Pearson's distance measures the strength and direction of a linear association between two data points based on their correlation coefficient. This method proved useful to assess the linear association between protein expression patterns across progressing precancerous lesions (Fig. 3E and F). In leukoplakia, COL6A2, FGB, VIM, Myosin-7 (MYH7) and COL6A1 exhibited strong positive correlations while CRNN and ANXA7 exhibited strong negative correlations. In erythroplakia, COL6A2, ANXA2, PRELP, and Tubulin beta-4B chain (TUBB4B) showed positive linear expression patterns.

3.4. Integrative approach for differential expression analysis

The intersection of these three statistical analysis methods identified proteins with linear varying expression across the different histological grades of leukoplakia and erythroplakia, as represented as a Venn diagram (Fig. 4). This integrative approach ensured the robust selection of identified proteins that were significantly altered during the progression of oral precursor lesions, and warrant further functional studies to elucidate their roles in the molecular mechanisms underlying oral cancer pathogenesis. Proteins with a linear varying expression across the histopathological phenotypes from hyperplasia, through dysplasia, to neoplasia are briefly discussed below and illustrated in Fig. 5.

3.5. Collagen type VI alpha 2 chain (COL6A2)

COL6A2 gene encodes the alpha (α)2 chain of type VI collagen, which is integral to the formation of the characteristic beaded microfilaments network in the extracellular matrix (ECM) [28]. COL6A2 plays a crucial role in cancer cell proliferation, migration, and invasion by modulating MAPK and Akt signaling pathways, inducing cell cycle arrest, and regulating the activity of MMP-2 and MMP-9 [29]. Co-expression analysis of COL6A members in pan-cancer has shown that COL6A1, COL6A2, and COL6A3 are highly positively correlated and highly expressed, suggesting a potential role in tumorigenesis [30]. Furthermore, significant and consistent overexpression of COL6A1, COL6A2, and COL6A3 genes has been reported in higher-grade gliomas, where they promote proliferation of glioma cells and inhibit their apoptosis indicating a potential correlation with tumor progression [31,32]. These findings collectively suggest that COL6A2 is correlated with the cell survival and cancer progression.

3.6. Annexin A2 (ANXA2)

Annexin A2, a calcium-dependent, phospholipid-binding protein, is found on various cell types and plays multiple roles in cancer progression. (i) On cell membranes, the ANXA2-S100A10 heterotetramer activates MMPs, increasing invasive ability [33]; (ii) In the cytoplasm, phosphorylated ANXA2 binds to actin filaments, enhancing migration [33], (iii) Within the nucleus, ANXA2 promotes epithelial-mesenchymal-transition EMT- and cancer stem cell-related transcription factors such as Snail, Twist, Oct4, Sox2, and Nanog [33], (iv) On the cell surface, ANXA2 facilitates the conversion of plasminogen to plasmin, which activates metalloproteinases that degrade extracellular matrix components, thereby promoting metastatic progression [34]. ANXA2 levels are elevated in OSCC tissues compared to adjacent normal oral epithelial tissues and OSCC cell lines [35,36]. Also, functional studies demonstrated that knockdown of ANXA2 in OSCC cell lines inhibited their proliferation, migration, and invasion capabilities [37]. With respect to signaling pathways, ANXA2-mediates activation of cell cycle regulation, epithelial-mesenchymal transition, and angiogenesis, such as the p53, c-Myc, STAT3, and ERK1/2 pathways [38–41]. ANXA2 therefore appears to be an important event in OSCC progression and is associated with adverse histopathological features.

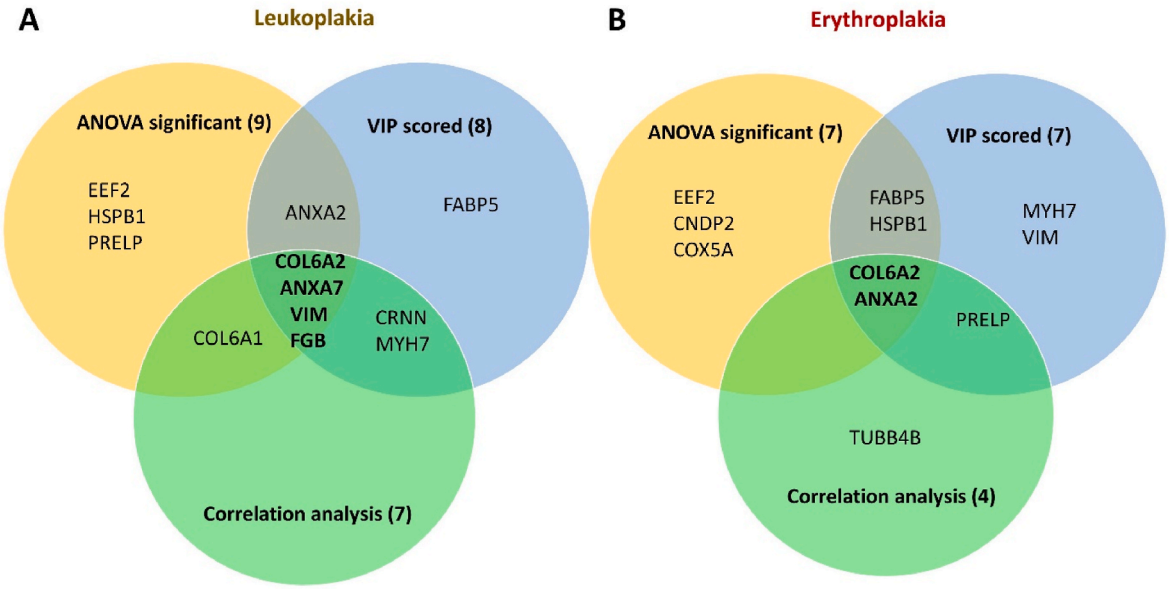

Fig. 4. Venn diagrams depicting the differentially expressed proteins associated with the progression of (A) leukoplakia, and (B) erythroplakia towards malignant transformation. This integrated statistical approach utilized the following criteria: One-way ANOVA with p-value <0.05 to identify proteins with significant mean differences between two or more groups. PLS-DA VIP scores >1.5 to select the most important proteins for discriminating the sample groups. Pearson correlation analysis to identify proteins based upon correlation coefficient (r): $0.6 \leq r \leq 1$ (positive correlation) and $-0.6 \leq r \leq 1$ (negative correlation).

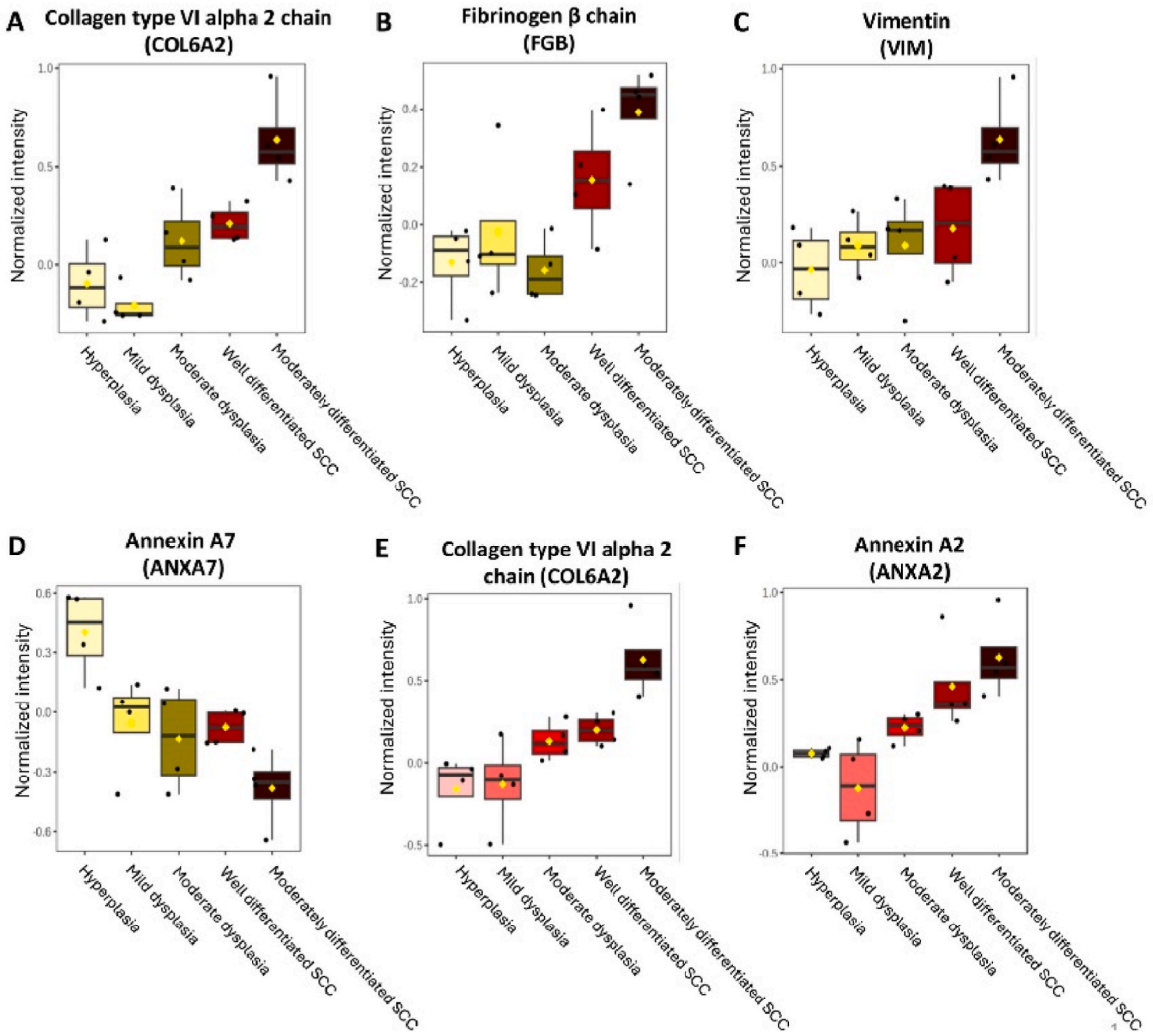

Fig. 5. Expression patterns of proteins with linearly varying expression across the histopathological grades of pre-cancerous lesions. (A–D) Leukoplakia, and (E–F) Erythroplakia. Corrected reporter intensities derived from MaxQuant analysis were subjected to median normalization followed by $\log_{10}$ transformation using MetaboAnalyst (v6.0).

3.7. Annexin A7 (ANXA7)

ANXA7 expression was found to decrease significantly with worsening histopathology in leukoplakia, from hyperplasia to more advanced stages. Annexin A7 (ANXA7), a calcium-dependent membrane fusion protein, plays a crucial role in calcium signaling pathways by interacting with the IP3 receptor and the PI3K/mTOR pathway, which are vital for regulating cell proliferation and survival, with disruptions potentially leading to uncontrolled cell growth and cancer [42]. The expression of annexin A7 (ANXA7) is selectively reduced in certain malignant tumors, and this down regulation correlated progression of the cancer [43,44].

3.8. Fibrinogen β chain (FGB)

FGB gene encodes the beta chain of fibrinogen, which is involved in coagulation and inflammation. In cancer, these processes can promote tumor growth and metastasis by creating a favorable microenvironment [45]. Interactions between the fibrinogen β chain and endothelial cells increase vascular permeability, allowing fibrinogen to leak into the tumor stroma, thereby supporting tumor growth and metastasis by enhancing nutrient and oxygen supply [46]. FGB plays crucial roles in wound healing, hemostasis, and tumor-related processes. FGB has an increased expression in cancers of bladder, larynx, lung, colon, and rectum as compared to the normal tissues [47,48]. These findings place the role of FGB in cancer progression in the right perspective.

3.9. Vimentin (VIM)

Vimentin, a type III intermediate filament protein in mesenchymal cells, is known to maintain cellular integrity and provide resistance against stress [49]. It is a key marker for EMT, in which epithelial cells gain mesenchymal traits that enhance their migration and invasion, critical for cancer metastasis [49]. In head and neck squamous cell carcinoma patients, EMT-related changes, characterized by loss of E-cadherin and increased vimentin expression, correlate with tumor progression and an increase of metastases [50]. Elevated Vimentin expression correlated with both poor cell differentiation and lymph node metastasis in tongue squamous cell carcinoma [51]. In keratinocytes isolated from carcinoma, vimentin protein expression increased compared to dysplasia and hyperplasia, with concomitant elevation in mRNA levels, suggesting transcriptional regulation [52]. These studies underscore vimentin's role in cancer biology, particularly in facilitating malignant transformation.

3.10. Deciphering progression from normal cell to cancer using functional enrichment

Functional enrichment with respect to the five proteins that are discussed above, in hyperplasia and cancer comprises of Cellular component, Biological process, Molecular function, and Reactome pathways (Table 2). The most notable 'Functional annotations' ($p < 0.0005$) are: (a) extracellular exosome; (b) collagen-containing extracellular matrix, (c) Hemostasis, (d) platelet aggregation, and (e) cell adhesion molecule binding. Exosomes are known to facilitate intercellular communication within the tumor micro-environment, promoting tumor growth [53]. OSCC-derived exosomes promote tumor growth and metastasis by activating signaling pathways like PI3K/Akt and MAPK/ERK through the secretion and uptake [54]. Collagen-containing extracellular matrix features as a cellular component in cancer for four of the five proteins. Induction of collagen cross-linking stiffens the ECM, promoting focal adhesions, enhances PI3 kinase activity, and induces the invasion of an oncogene-initiated epithelium facilitating tumor progression [55]. These functional processes, either individually or in unison, are responsible for the transformation of normal cells to neoplasia through the stages of dysplasia.

3.11. Limitations

(1) Cellular heterogeneity in protein expression exists even within the same phenotype due to patient-specific biological differences, timing of biopsy collection, influenced by individual health conditions. Also, variations in biopsy depth and width can impact the proteomic profile, given that proteins from different epithelial layers contribute distinctively. Absence of follow-up samples for successive or progressive phenotypes was therefore a limitration. Our analysis therefore provides snap shots across individual phenotypes that depicts the dynamic process that exists in the progression that is involved in the transition from precancerous lesions to overt oral squamous cell carcinoma. (2) Our study utilized biopsy samples, particularly premalignant specimens, which were available in limited quantities. This scarcity posed a major challenge in conducting additional validation experiments.

4. Conclusion

Label based proteomics is an ideal platform to study oral cancer progression. Collagen type VI alpha 2 chain, Fibrinogen β chain, and Vimentin were found to have increased linear expression across precancer phenotypes of leukoplakia to cancer, while Annexin A7 was seen to be having a linear decreasing expression. Collagen type VI alpha 2 chain and Annexin A2 had increased linear expression across precancer phenotypes of erythroplakia to cancer. These differentially expressed proteins mediate cancer progression mainly through extracellular exosome collagen-containing extracellular matrix, hemostasis, platelet aggregation, and cell adhesion molecule binding. The study has translational value for early detection, risk stratification, and potential therapeutic targeting of oral premalignant lesions and in its prevention to malignant forms.

CRediT authorship contribution statement

Vipra Sharma: Writing – original draft, Methodology, Formal analysis. **Sundararajan Baskar Singh:** Project administration, Funding acquisition, Conceptualization. **Sabyasachi Bandyopadhyay:** Validation, Software, Methodology. **Kapil Sikka:** Supervision, Resources,

Table 2
Functional annotation of identified proteins.

	Cellular Component	Biological Process		Molecular Function	Reactome pathways	
	Cancer	Hyperplasia	Cancer	Cancer	Hyperplasia	Cancer
COL6A2	- Extracellular exosome (GO:0070062) - Collagen-containing extracellular matrix (GO:0062023)				- ECM proteoglycans (R-HSA:300017)	- Axon guidance (R-HSA:422475)
ANXA2	- Extracellular exosome (GO:0070062) - Collagen-containing extracellular matrix (GO:0062023) - Vacuolar lumen (GO:0005775) - Melanosome (GO:0042470)	- Hemostasis (GO:0007599)	- Cell adhesion mediator activity (GO:0098631) - Phospholipase inhibitor activity (GO:0004859) - Hemostasis (GO:0007599)	- Cell adhesion molecule binding (GO:0050839)		- Interleukin-12 signaling (R-HAS:902059) - Neutrophil degranulation (R-HSA:679869)
ANXA7	- Extracellular exosome (GO:0070062) - Collagen-containing extracellular matrix (GO:0062023)			- Cell adhesion molecule binding (GO:0050839)		
FGB	- Extracellular exosome (GO:0070062) - Collagen-containing extracellular matrix (GO:0062023)	- Hemostasis (GO:0007599) - Homotypic cell-cell adhesion (GO:0034109) - Platelet aggregation (GO:0070527)	- Hemostasis (GO:0007599) - Homotypic cell-cell adhesion (GO:0034109) - Platelet aggregation (GO:0070527) - Antimicrobial humoral response (GO:0019730)	- Cell adhesion molecule binding (GO:0050839)		
VIM	- Extracellular exosome (GO:0070062) - Focal adhesion (GO:0005925)				- Apoptotic cleavage of cellular proteins (R-HSA:111465)	- Programmed cell death (R-HSA:535780) - Chaperone mediated autophagy (R-HSA:961382)

Methodology. **Aanchal Kakkar:** Supervision, Resources, Methodology. **Gururao Hariprasad:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

Authors report no conflict of interest in this work.

Acknowledgements

Funding from DST (EEQ/2018/001412), Governement of India is acknowledged.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchanger Consortium via the PRIDE partner repository with the data set identifier PXD054190.

References

[1] J. Ferlay, M. Ervik, F. Lam, M. Laversanne, M. Colombet, L. Mery, M. Piñeros, A. Znaor, I. Soerjomataram, F. Bray, Global Cancer Observatory: Cancer Today, International Agency for Research on Cancer, Lyon, France, 2024. Available from: https://gco.iarc.who.int/today. (Accessed 28 March 2024).

[2] F. Bray, M. Laversanne, H. Sung, J. Ferlay, R.L. Siegel, I. Soerjomataram, A. Jemal, Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 74 (2024) 229–263, https://doi.org/10.3322/caac.21834.

[3] S.-H. Chen, S.-Y. Hsiao, K.-Y. Chang, J.-Y. Chang, New insights into oral squamous cell carcinoma: from clinical aspects to molecular tumorigenesis, Int. J. Mol. Sci. 22 (2021), https://doi.org/10.3390/ijms22052252.

[4] O. Iocca, T.P. Sollecito, F. Alawi, G.S. Weinstein, J.G. Newman, A. De Virgilio, P. Di Maio, G. Spriano, S. Pardiñas López, R.M. Shanti, 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 (2020) 539–555, https://doi.org/10.1002/hed.26006.

[5] J. Reibel, N. Gale, J. Hille, J. Hunt, M. Lingen, S. Muller, P. Sloan, W. Tilakaratne, W. Westra, M. Williams, N. Vigneswaran, H. Fatani, E. Odell, R. Zain, Oral potentially malignant disorders and oral epithelial dysplasia, in: WHO Classification of Head and Neck Tumours, 2017.

[6] S. Warnakulasuriya, O. Kujan, J.M. Aguirre-Urizar, J. V Bagan, M.Á. González-Moles, A.R. Kerr, G. Lodi, F.W. Mello, L. Monteiro, G.R. Ogden, P. Sloan, N. W. Johnson, 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 Dis. 27 (2021) 1862–1880, https://doi.org/10.1111/odi.13704.

[7] C. Zhang, B. Li, X. Zeng, X. Hu, H. Hua, The global prevalence of oral leukoplakia: a systematic review and meta-analysis from 1996 to 2022, BMC Oral Health 23 (2023) 645, https://doi.org/10.1186/s12903-023-03342-y.

[8] J.M. Aguirre-Urizar, I. Lafuente-Ibáñez de Mendoza, S. Warnakulasuriya, Malignant transformation of oral leukoplakia: systematic review and meta-analysis of the last 5 years, Oral Dis. 27 (2021) 1881–1895, https://doi.org/10.1111/odi.13810.

[9] A.I. Lorenzo-Pouso, I. Lafuente-Ibáñez de Mendoza, M. Pérez-Sayáns, A. Pérez-Jardón, C.M. Chamorro-Petronacci, A. Blanco-Carrión, J.M. Aguirre-Urízar, Critical update, systematic review, and meta-analysis of oral erythroplakia as an oral potentially malignant disorder, J. Oral Pathol. Med. 51 (2022) 585–593, https://doi.org/10.1111/jop.13304.

[10] M. Srinivasan, S.D. Jewell, Evaluation of TGF-α and EGFR expression in oral leukoplakia and oral submucous fibrosis by quantitative immunohistochemistry, Oncology 61 (2001) 284–292, https://doi.org/10.1159/000055335.

[11] J.R. Grandis, S.D. Drenning, Q. Zeng, S.C. Watkins, M.F. Melhem, S. Endo, D. E. Johnson, L. Huang, Y. He, J.D. Kim, Constitutive activation of Stat3 signaling abrogates apoptosis in squamous cell carcinogenesis in vivo, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 4227–4232, https://doi.org/10.1073/pnas.97.8.4227.

[12] J. Rubin Grandis, D.J. Tweardy, M.F. Melhem, Asynchronous modulation of transforming growth factor α and epidermal growth factor receptor protein expression in progression of premalignant lesions to head and neck squamous cell carcinoma, Clin. Cancer Res. 4 (1998).

[13] S. Soni, J. Kaur, A. Kumar, N. Chakravarti, M. Mathur, S. Bahadur, N.K. Shukla, S. V.S. Deo, R. Ralhan, Alterations of rb pathway components are frequent events in patients with oral epithelial dysplasia and predict clinical outcome in patients with squamous cell carcinoma, Oncology 68 (2005) 314–325, https://doi.org/10.1159/000086970.

[14] H. Xiao, A. Langerman, Y. Zhang, O. Khalid, S. Hu, C.-X. Cao, M.W. Lingen, D.T. W. Wong, Quantitative proteomic analysis of microdissected oral epithelium for cancer biomarker discovery, Oral Oncol. 51 (2015) 1011–1019, https://doi.org/10.1016/j.oraloncology.2015.08.008.

[15] R. Ralhan, L.V. Desouza, A. Matta, S.C. Tripathi, S. Ghanny, S. Dattagupta, A. Thakar, S.S. Chauhan, K.W. Michael Siu, ITRAQ-multidimensional liquid chromatography and tandem mass spectrometry-based identification of potential biomarkers of oral epithelial dysplasia and novel networks between inflammation and premalignancy, J. Proteome Res. 8 (2009), https://doi.org/10.1021/pr800501j.

[16] V. Mohanty, Y. Subbannayya, S. Patil, V.N. Puttamallesh, MohdA. Najar, K. K. Datta, S.M. Pinto, S. Begum, N. Mohanty, S. Routray, R. Abdulla, J.G. Ray, D. Sidransky, H. Gowda, T.S.K. Prasad, A. Chatterjee, Molecular alterations in oral cancer using high-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue, J. Cell Commun. Signal 15 (2021) 447–459, https://doi.org/10.1007/s12079-021-00609-3.

[17] V. Sharma, S. Bandyopadhyay, K. Sikka, A. Kakkar, G. Hariprasad, S.B. Singh, Label-free proteomics of oral mucosa tissue to identify potential biomarkers that can flag predilection of precancerous lesions to oral cell carcinoma: a preliminary study, Dis. Markers 2023 (2023) 1–16, https://doi.org/10.1155/2023/1329061.

[18] V. Sharma, M.V. Rajan, S.B. Singh, S. Bandyopadhyay, K. Sikka, A. Kakkar, G. Hariprasad, Comparative proteomics of oral squamous cell carcinoma tissue in consumers and non-consumers of tobacco, J. Protein Proteonomics (2024), https://doi.org/10.1007/s42485-024-00151-x.

[19] S.-H. Yu, P. Kyriakidou, J. Cox, Isobaric matching between runs and novel PSM-level normalization in MaxQuant strongly improve reporter ion-based quantification, J. Proteome Res. 19 (2020) 3945–3954, https://doi.org/10.1021/acs.jproteome.0c00209.

[20] E.W. Deutsch, A. Csordas, Z. Sun, A. Jarnuczak, Y. Perez-Riverol, T. Ternent, D. S. Campbell, M. Bernal-Llinares, S. Okuda, S. Kawano, R.L. Moritz, J.J. Carver, M. Wang, Y. Ishihama, N. Bandeira, H. Hermjakob, J.A. Vizcaíno, The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition, Nucleic Acids Res. 45 (2017) D1100–D1106, https://doi.org/10.1093/nar/gkw936.

[21] Y. Perez-Riverol, J. Bai, C. Bandla, D. García-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D.J. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher, M. Walzer, S. Wang, A. Brazma, J.A. Vizcaíno, The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences, Nucleic Acids Res. 50 (2022) D543–D552, https://doi.org/10.1093/nar/gkab1038.

[22] Z. Pang, Y. Lu, G. Zhou, F. Hui, L. Xu, C. Viau, A.F. Spigelman, P.E. MacDonald, D. S. Wishart, S. Li, J. Xia, MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation, Nucleic Acids Res. (2024), https://doi.org/10.1093/nar/gkae253.

[23] G. Bindea, B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W. H. Fridman, F. Pagès, Z. Trajanoski, J. Galon, ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks, Bioinformatics 25 (2009), https://doi.org/10.1093/bioinformatics/btp101.

[24] S. Roy, S. Girotra, A. Radhakrishnan, S. Basu, Prevalence and determinants of tobacco consumption and oral cancer screening among men in India: evidence from a nationally representative cross-sectional Survey, J. Public Health (2023), https://doi.org/10.1007/s10389-023-02161-3.

[25] S. Farooqui, S. Mohammad, D. Mehrotra, A.A. Mahdi, S. Bhattacharya, G.G. Agarwal, S. Srivastava, Study on prevalence and sociocultural aspects of tobacco use in India., Natl. J. Maxillofac. Surg. 10 (n.d.) 182–190. https://doi.org/10.4103/njms.NJMS_82_18.

[26] R. Thakur, A. Thakar, R.K. Malhotra, A. Sharma, A. Kakkar, Tumor-host interface in oral squamous cell carcinoma: impact on nodal metastasis and prognosis, Eur. Arch. Oto-Rhino-Laryngol. 278 (2021) 5029–5039, https://doi.org/10.1007/s00405-021-06756-y.

[27] N. Santosh, K.K. McNamara, F.M. Beck, J.R. Kalmar, Expression of cornulin in oral premalignant lesions, Oral. Surg. Oral Med. Oral Pathol. Oral Radiol. 127 (2019) 526–534, https://doi.org/10.1016/j.oooo.2019.02.003.

[28] M. Cescon, F. Gattazzo, P. Chen, P. Bonaldo, Collagen VI at a glance, J. Cell Sci. (2015), https://doi.org/10.1242/jcs.169748.

[29] X.-M. Piao, B. Hwang, P. Jeong, Y. Byun, H. Kang, S. Seo, W. Kim, J.-Y. Lee, Y.-S. Ha, Y.-S. Lee, I. Kim, Y. Choi, E.-J. Cha, S.-K. Moon, S. Yun, W.-J. Kim, Collagen type VI-α1 and 2 repress the proliferation, migration and invasion of bladder cancer cells, Int. J. Oncol. 59 (2021) 37, https://doi.org/10.3892/ijo.2021.5217.

[30] X. Li, Z. Li, S. Gu, X. Zhao, A pan-cancer analysis of collagen VI family on prognosis, tumor microenvironment, and its potential therapeutic effect, BMC Bioinf. 23 (2022) 390, https://doi.org/10.1186/s12859-022-04951-0.

[31] M. Cescon, E. Rampazzo, S. Bresolin, F. Da Ros, L. Manfreda, A. Cani, A. Della Puppa, P. Braghetta, P. Bonaldo, L. Persano, Collagen VI sustains cell stemness and chemotherapy resistance in glioblastoma, Cell. Mol. Life Sci. 80 (2023) 233, https://doi.org/10.1007/s00018-023-04887-5.

[32] X. Hong, J. Zhang, J. Zou, J. Ouyang, B. Xiao, P. Wang, X. Peng, Role of COL6A2 in malignant progression and temozolomide resistance of glioma, Cell. Signal. 102 (2023) 110560, https://doi.org/10.1016/j.cellsig.2022.110560.

[33] C.-Y. Chen, Y.-S. Lin, C.-H. Chen, Y.-J. Chen, Annexin A2-mediated cancer progression and therapeutic resistance in nasopharyngeal carcinoma, J. Biomed. Sci. 25 (2018) 30, https://doi.org/10.1186/s12929-018-0430-8.

[34] N.A. Lokman, M.P. Ween, M.K. Oehler, C. Ricciardelli, The role of annexin A2 in tumorigenesis and cancer progression, Cancer Microenviron. 4 (2011) 199–208, https://doi.org/10.1007/s12307-011-0064-9.

[35] L. Zhong, K. Wei, X. Yang, L. Zhang, X. Zhou, H. Pan, J. Li, W. Chen, Z. Zhang, Increased expression of Annexin A2 in oral squamous cell carcinoma, Arch. Oral Biol. 54 (2009) 17–25, https://doi.org/10.1016/j.archoralbio.2008.08.006.

[36] Y. Ma, H. Wang, Clinical significance of Annexin A2 expression in oral squamous cell carcinoma and its influence on cell proliferation, migration and invasion, Sci. Rep. 11 (2021) 5033, https://doi.org/10.1038/s41598-021-84675-y.
[37] H.-J. Zhang, Annexin A2 silencing inhibits invasion, migration, and tumorigenic potential of hepatoma cells, World J. Gastroenterol. 19 (2013) 3792, https://doi.org/10.3748/wjg.v19.i24.3792.
[38] C.-Y. Wang, C.-L. Chen, Y.-L. Tseng, Y.-T. Fang, Y.-S. Lin, W.-C. Su, C.-C. Chen, K.-C. Chang, Y.-C. Wang, C.-F. Lin, Annexin A2 silencing induces G2 arrest of non-small cell lung cancer cells through p53-dependent and -independent mechanisms, J. Biol. Chem. 287 (2012) 32512–32524, https://doi.org/10.1074/jbc.M112.351957.
[39] S. Ma, C.-C. Lu, L.-Y. Yang, J.-J. Wang, B.-S. Wang, H.-Q. Cai, J.-J. Hao, X. Xu, Y. Cai, Y. Zhang, M.-R. Wang, ANXA2 promotes esophageal cancer progression by activating MYC-HIF1A-VEGF axis, J. Exp. Clin. Cancer Res. 37 (2018) 183, https://doi.org/10.1186/s13046-018-0851-y.
[40] J. Yuan, Y. Yang, Z. Gao, Z. Wang, W. Ji, W. Song, F. Zhang, R. Niu, Tyr23 phosphorylation of Anxa2 enhances STAT3 activation and promotes proliferation and invasion of breast cancer cells, Breast Cancer Res. Treat. 164 (2017) 327–340, https://doi.org/10.1007/s10549-017-4271-z.
[41] B. Wu, F. Zhang, M. Yu, P. Zhao, W. Ji, H. Zhang, J. Han, R. Niu, Up-regulation of Anxa2 gene promotes proliferation and invasion of breast cancer MCF-7 cells, Cell Prolif. 45 (2012) 189–198, https://doi.org/10.1111/j.1365-2184.2012.00820.x.
[42] M. Srivastava, A. Bera, O. Eidelman, M.B. Tran, C. Jozwik, M. Glasman, X. Leighton, H. Caohuy, H.B. Pollard, A dominant-negative mutant of ANXA7 impairs calcium signaling and enhances the proliferation of prostate cancer cells by downregulating the IP3 receptor and the PI3K/mTOR pathway, Int. J. Mol. Sci. 24 (2023) 8818, https://doi.org/10.3390/ijms24108818.
[43] K.S. Hung, S.L. Howng, Prognostic significance of annexin VII expression in glioblastomas multiforme in humans, J. Neurosurg. 99 (2003), https://doi.org/10.3171/jns.2003.99.5.0886.
[44] P. Hsu, M. Huang, H. Chen, P. Hsu, T. Lai, J. Wang, G. Lo, K. Lai, C. Tseng, M. Hsiao, The significance of ANXA7 expression and its correlation with poor cellular differentiation and enhanced metastatic potential of gastric cancer, J. Surg. Oncol. 97 (2008) 609–614, https://doi.org/10.1002/jso.21046.
[45] O. Repetto, S. Maiero, R. Magris, G. Miolo, M. Cozzi, A. Steffan, V. Canzonieri, R. Cannizzaro, V. De Re, Quantitative proteomic approach targeted to fibrinogen β chain in tissue gastric carcinoma, Int. J. Mol. Sci. 19 (2018) 759, https://doi.org/10.3390/ijms19030759.
[46] X. Wu, X. Yu, C. Chen, C. Chen, Y. Wang, D. Su, L. Zhu, Fibrinogen and tumors, Front. Oncol. 14 (2024), https://doi.org/10.3389/fonc.2024.1393599.
[47] M. Lindén, S.B. Lind, C. Mayrhofer, U. Segersten, K. Wester, Y. Lyutvinskiy, R. Zubarev, P. Malmström, U. Pettersson, Proteomic analysis of urinary biomarker candidates for nonmuscle invasive bladder cancer, Proteomics 12 (2012) 135–144, https://doi.org/10.1002/pmic.201000810.
[48] C. Zhang, W. Leng, C. Sun, T. Lu, Z. Chen, X. Men, Y. Wang, G. Wang, B. Zhen, J. Qin, Urine proteome profiling predicts lung cancer from control cases and other tumors, EBioMedicine 30 (2018) 120–128, https://doi.org/10.1016/j.ebiom.2018.03.009.
[49] A. Satelli, S. Li, Vimentin in cancer and its potential as a molecular target for cancer therapy, Cell. Mol. Life Sci. 68 (2011) 3033–3046, https://doi.org/10.1007/s00018-011-0735-1.
[50] M.M. Nijkamp, P.N. Span, I.J. Hoogsteen, A.J. van der Kogel, J.H.A.M. Kaanders, J. Bussink, Expression of E-cadherin and vimentin correlates with metastasis formation in head and neck squamous cell carcinoma patients, Radiother. Oncol. 99 (2011) 344–348, https://doi.org/10.1016/j.radonc.2011.05.066.
[51] P.-F. Liu, B.-H. Kang, Y.-M. Wu, J.-H. Sun, L.-M. Yen, T.-Y. Fu, Y.-C. Lin, H.-H. Liou, Y.-S. Lin, H.-C. Sie, I.-C. Hsieh, Y.-K. Tseng, C.-W. Shu, Y.-D. Hsieh, L.-P. Ger, Vimentin is a potential prognostic factor for tongue squamous cell carcinoma among five epithelial–mesenchymal transition-related proteins, PLoS One 12 (2017) e0178581, https://doi.org/10.1371/journal.pone.0178581.
[52] S. Sawant, M. Vaidya, D. Chaukar, H. Alam, C. Dmello, P. Gangadaran, S. Kannan, S. Kane, P. Dange, N. Dey, K. Ranganathan, A. D'Cruz, Clinical significance of aberrant vimentin expression in oral premalignant lesions and carcinomas, Oral Dis. 20 (2014) 453–465, https://doi.org/10.1111/odi.12151.
[53] Y. Lu, Z. Zheng, Y. Yuan, J.L. Pathak, X. Yang, L. Wang, Z. Ye, W.C. Cho, M. Zeng, L. Wu, The emerging role of exosomes in oral squamous cell carcinoma, Front. Cell Dev. Biol. 9 (2021), https://doi.org/10.3389/fcell.2021.628103.
[54] S. Sento, E. Sasabe, T. Yamamoto, Application of a persistent heparin treatment inhibits the malignant potential of oral squamous carcinoma cells induced by tumor cell-derived exosomes, PLoS One 11 (2016), https://doi.org/10.1371/journal.pone.0148454.
[55] K.R. Levental, H. Yu, L. Kass, J.N. Lakins, M. Egeblad, J.T. Erler, S.F.T. Fong, K. Csiszar, A. Giaccia, W. Weninger, M. Yamauchi, D.L. Gasser, V.M. Weaver, Matrix crosslinking forces tumor progression by enhancing integrin signaling, Cell 139 (2009) 891–906, https://doi.org/10.1016/j.cell.2009.10.027.