

RESEARCH

Open Access



The spatial distribution of intermediate fibroblasts and myeloid-derived cells dictate lymph node metastasis dynamics in oral cancer

Soni Shaikh^{1,2†}, Harsh Dhar^{3†}, Manju Moorthy⁴, Vijayalakshmi Bhat⁴, Sangramjit Basu⁵, Devmalya Banerjee⁶, Deepak Kumar Mishra¹, Sourav Datta^{3*} and Geetashree Mukherjee^{1*} 

Abstract

Background Oral cancer poses a significant health challenge due to limited treatment protocols and therapeutic targets. We aimed to investigate the invasive margins of gingivo-buccal oral squamous cell carcinoma (GB-OSCC) tumors in terms of the localization of genes and cell types within the margins at various distances that could lead to nodal metastasis.

Methods We collected tumor tissues from 23 resected GB-OSCC samples for gene expression profiling using digital spatial transcriptomics. We monitored differential gene expression at varying distances between the tumor and its microenvironment (TME), and performed a deconvolution study and immunohistochemistry to identify the cells and genes regulating the TME.

Results We found that the tumor–stromal interface (a distance up to 200 μ m between tumor and immune cells) is the most active region for disease progression in GB-OSCC. The most differentially expressed apex genes, such as FN1 and COL5A1, were located at the stromal ends of the margins, and together with enrichment of the extracellular matrix (ECM) and an immune-suppressed microenvironment, were associated with lymph node metastasis. Intermediate fibroblasts, myocytes, and neutrophils were enriched at the tumor ends, while cancer-associated fibroblasts (CAFs) were enriched at the stromal ends. The intermediate fibroblasts transformed into CAFs and relocated to the adjacent stromal ends where they participated in FN1-mediated ECM modulation.

Conclusion We have generated a functional organization of the tumor–stromal interface in GB-OSCC and identified spatially located genes that contribute to nodal metastasis and disease progression. Our dataset might now be mined to discover suitable molecular targets in oral cancer.

Keywords Extracellular matrix, Cancer-associated fibroblasts, Gingivo-buccal oral squamous cell carcinoma, Spatial transcriptomics, Tumor margin

[†]Soni Shaikh and Harsh Dhar equally contributed to this work.

*Correspondence:

Sourav Datta

sourav.datta@medicasynergie.in

Geetashree Mukherjee

geetashree.mukherjee@tmckolkata.com

Full list of author information is available at the end of the article



Introduction

Oral cancer represents a significant global health challenge with an annual new case-to-death ratio of 2:1 (GLOBOCAN, 2022). Two-thirds of cases are diagnosed in developing countries, with the highest incidence in South-Central Asia. Indeed, in India, oral cancer is the most diagnosed cancer among males, comprising 16.1% of all male cancer cases. The majority (40%) of oral cancers occur in the gingivobuccal region as squamous cell carcinomas (GB-OSCC) [1, 2]. Unfortunately, treatment protocols are limited to surgery, radiotherapy and chemotherapy, as no viable targeted therapy currently exists. As such, the prognosis of the disease remains poor: the 5-year disease-free survival (DFS) ranges between 50 and 60%, but in locally advanced cases (often driven by lymph-node metastasis) it declines to 12% [3, 4]. There is thus an urgent need to explore new directions for therapeutic strategies.

The tumor microenvironment (TME) has a critical role in mediating tumor metastasis and tumor aggressiveness. The most notable components of the TME include non-neoplastic cells, cytokines and chemokines, the blood and lymphatic vascular systems, the extracellular matrix (ECM) and its associated growth factors, innate and adaptive immune cells, cancer-associated fibroblasts (CAFs), and some cancer stem cells (CSCs) [5]. Although proteins related to the ECM are produced by stromal and tumor cells, CAFs are the main source of these proteins [6].

Cancer cell–TME interactions at the tumor–stroma interface control the signaling dynamics that regulate immune-cell functions, cell migration, and metastasis [7]. These interactions are primarily regulated via integrin receptors that directly bind to ECM proteins such as collagens and fibronectin subtypes [5, 8, 9]. Alongside, tumor-associated macrophages (TAMs) secrete chemokines in the TME to maintain an immune-suppressive microenvironment [10]. To suppress immune surveillance in the TME, tumor cells reversibly alter the frontline cell phenotypes via cell–cell and cell–matrix contact, cytokine secretion, and exosome release [11, 12].

The cell types in the TME are not equally distributed. For example, tumor margins more typically exhibit a high density of immune cells compared to distal regions of the TME. Such unequal cell distribution implies a differential spatial distribution of gene expression and active signaling would exist in defined regions [13, 14]. These non-linear characteristics in gene and protein expression has rendered the TME an attractive target for cancer genomics and therapeutic studies in recent years. Such studies in the context of oral cancer, however, are limited. One study has shown the differential gene expression between tumor centers and invasive margins in oral cancer [15].

The study demonstrated a conserved genomic signature at the tumor margins. We aimed to investigate the spatial cell and gene expression distribution in the TME of 23 samples of GB-OSCC, and to infer the mechano-transduction machinery driving cancer-cell invasion and disease progression by spatial transcriptomics.

Materials and methods

Ethics and study approval

Written consent was obtained from all patients included in this study. The study was approved by the Institutional Review Board of Tata Medical Center, Kolkata, India (Ref. no. EC/GOVT/23/17) and conducted in accordance with the Helsinki Declaration.

Study cohort

Tumor tissues were collected from resected specimens of HPV-negative GB-OSCC (91% moderately differentiated) taken during primary surgical procedures from 23 patients. 21 patients were males, with 57% > 50 years-of-age (Table S1). A total of 10 cases had lymph node metastases on diagnosis. Only patients > 18 years with newly diagnosed GB-OSCC were included; patients who had received prior treatment were excluded.

Study design

The overall GeoMx study design is presented in (Figure S1) and work flow of current study is presented in Figure S2. Accordingly, the accepted quality of data of 87AOI within 48 ROI is presented in Table S2. Table S3 represents a complete annotation of each profiled AOI and the distance between tumor and immune cells.

Tissue samples for Nanostring and TMA construction

Tumor tissues were fixed in formalin and then embedded in paraffin (FFPE) to construct a tissue microarray (TMA); diameter 3 mm using a manual tissue array produced by UNITMA (Model UT06, Seoul, Korea) as per the standard guidelines [16]. The punches from donor blocks included the invasive edge and tumor center [17] and arrayed on TMA blocks. Four punches containing normal squamous epithelium with normal stroma were included.

Slide preparation for Nanostring digital spatial transcriptomics and sequencing

Slides were prepared according to the GeoMx DSP instructions. Briefly, 5 μ m FFPE tissue sections of the TMA blocks were mounted on a Superfrost® Plus Micro Slide (Cat: 1,255,015) and baked at 65 °C for 1 h. Then, the sections were immersed in xylene (three times for 5 min each) and washed in 100% ethanol (two times for 5 min each), followed by 95% ethanol, and then

1×PBS. Next, the slides were immersed in hot, nuclease-free water (10 s) before antigen retrieval using Antigen retrieval solution pH 9.0 (Invitrogen, Cat: 2450723) at 100 °C for 15 min. The slides were washed again in 1×PBS, before being incubated in proteinase K in 1×PBS (1 µg/ml) for 15 min at 37 °C, and finally washed with 1×PBS. After fixation, the tissues were washed with 10% neutral-buffered formalin (NBF) for 5 min, followed by two washes of 5 min each in NBF stop buffer (0.1 M Tris Base, 0.1 M Glycine), and one wash in 1×PBS. The slides were incubated overnight at 37 °C with 4 nM of each Human Cancer Transcriptome Atlas probe diluted in Buffer R (NanoString, Cat: HCTA22004), according to the instructions of the NanoStringGeoMx RNA-NGS slide preparation manual.

Slides were hybridized using HybriSlip hybridization covers (Grace BioLabs, Cat: 122085). The following day, the HybriSlip covers were gently removed by dipping the slides in saline-sodium citrate (SSC) buffer (Invitrogen & Cat: 1248542), and two stringent washes were performed for 25 min each in formamide and 4×SSC buffer at 37 °C. The tissues were then washed for 5 min in 2×SSC buffer and placed in a humid chamber for blocking with Buffer W (GeoMx RNA slide prep kit PCLN, Nanostring Technologies, Item code: 121300313) for 30 min at room temperature.

Next, the tissues were stained with a Solid Tumor (TME) morphology kit (Nanostring Technologies, Item code: 121300310) and SYTO 13 (GeoMx Nuclear stain morphology kit, Item no. 121300303) for 1 h at room temperature. The following immunofluorescent antibodies were used to stain and identify the tissue morphology in the processed slides: PanCK (pan-cytokeratin) to identify the epithelial cells, CD45 to identify immune cells, and SYTO 13 to label DNA. The slides were washed twice in fresh 2×SSC and loaded on the GeoMx Digital Spatial Profiler (DSP), and images were captured at 20× magnification.

Expert onco-pathologists identified and calculated the approximate distances between tumor and immune cells based on tissue morphology from the H&E images and using the scale bar of the GeoMx images as a guide (Fig. S3A, B). Regions of interest (ROI) were selected according to morphologic markers and previously stained H&E slides. Automatic segmentation of ROI based on panCK marker (Cy3 channel) was used to define areas of illumination (AOI), while other fluorescent channels were set to Ø(ignore), allowing the separation of the “tumor end” (panCK+) from the adjacent “stromal end” (panCK-) (Fig. S3B, C). The segment of the panCK+ area (AOI-1) at the tumor margins comprised tumor cells and host cells that were in touching. The segment of panCK- area

(AOI-2) contained only cells from the TME at various distances towards the normal end of the tissue sections.

Library preparation for NanostringGeoMx-NGS

A total of 48 ROIs were segmented into 87 AOI. Under UV light (385 nm), the indexed oligonucleotides were released and deposited into a 96-well plate through a microcapillary. The 96-well plates were dried on a PCR machine at 65 °C for 1 h and reconstituted in 10 µl DEPC-treated water (Ambion, cat: AM9922). Sequencing libraries were prepared according to the NanoString GeoMx-NGS Read out Library Prep manual. PCR was performed using NanoString Seq Code primers. PCR products were then pooled in equal volumes and purified with AMPure XP beads (Beckman Coulter, cat no. A63880) twice. Libraries were sequenced on an Illumina NovaSeq 6000, with 2×151 bp paired-end reads.

NanostringGeoMx CTA data analysis

The sequenced FASTQ files were processed and converted to a Digital Count Conversion (DCC) format using the GeoMx NGS Pipeline (v.2.3.3.10, NanoString) according to the following steps: (i) The adapters were trimmed using “trim galore”, (ii) the paired-end reads were merged using “flash2”, and (iii) the reads were aligned to the whitelist of RTSIDs (probe barcodes) corresponding to the target genes in the CTA panel and UMI extraction using “bowtie2”. UMI tools were then used to remove PCR duplicates and convert the read counts to digital ones. Individual ROI or AOI resulted in single DCC files. The DCC files were analyzed using “GeoMxTools” in R [18]. R version 4.2.0 was used throughout, unless stated otherwise. The analysis within the GeoMxTools pipeline included Quality control (QC) for both probes and segments, normalization, unsupervised clustering, and differential gene expression using the LMM model. The following parameters were used for the QC of the segments: min Segment Reads=1000, percent Trimmed=80, percent Stitched=80, percent Aligned=80, percent Saturation=50, min Negative Count=2, max NTC Count=1000, min Nuclei=60, min Area=5000 and min LOQ=2. The QC of probes (target genes) was based on the geometric mean of each probe count from all segments divided by the geometric mean of all probe counts from all segments (>0.1) and the percentage of outlier segments (Grubb’s test <20%).

Subsequently, segments and target gene counts were normalized using the Q3 method. Differential gene expression profiling was performed for the spatial ROIs from node-positive cases against the node-negative cases using the linear mixed model (LMM) package in R [19], with Bonferroni-Hochberg (BH) correction and random

intercept for tissue sub-sampling. Fold changes were considered significant for p -value < 0.05 .

The DAVID web server (<https://david.ncifcrf.gov/>) was used on the significant (p -value < 0.05) differentially expressed genes to determine the functional enrichment within the category of gene ontology (GO), biological processes, and molecular function, and REACTOME pathway enrichment.

The SpatialDecon algorithm was used to determine the immune cell type abundance for deconvolution within the ROI (safeTME profile matrix was applied) [20]. The same algorithm to deconvolute the abundance of OSCC-specific cell types was used, with a custom profile matrix generated from the scRNA-Seq data obtained from the GSE103322 dataset [21].

Gene set variation analysis

Gene set variation analysis was performed using the “GSVA” package in R [22], which profiled the enrichment of the MSigDB hallmark (<https://www.gsea-msigdb.org/gsea/msigdb/human/genesets.jsp?collection=H>) and OSCC metastasis-specific gene sets [21, 23], in the different ROIs, as follows: Tumor margin (distance between immune and tumor cells $< 200 \mu\text{m}$); tumor center (immune and tumor cells $< 200 \mu\text{m}$); immune and tumor cells $> 200 \mu\text{m}$; immune and tumor cells $> 300 \mu\text{m}$ and immune and tumor cells $> 500 \mu\text{m}$. The GSVA scores from the ROIs with a similar nodal status, segmentation status, ROI type, and distance between tumor and immune cells were averaged to visualize the trend of GSVA scores as a heatmap using the “pheatmap” package in R (<https://cran.r-project.org/web/packages/pheatmap/index.html>). The same method was used to score the hallmark, metastasis-specific, and CAF gene sets [21].

Protein–protein interaction (PPI) network of differentially expressed genes (DEGs)

The STRING database (v12.0, <http://string-db.org>) was used to predict the PPIs between DEGs (p -value < 0.05). A confidence score > 0.7 was considered significant for filtering the interactions from the STRING database. Cytoscape software (v3.10.0, <http://www.cytoscape.org/>) was used to construct the PPI network. The Molecular Complex Detection (MCODE) app (with “Degree cut-off = 2”, “node score cutoff = 0.2”, “k-core = 2” and “max. depth = 100” cut-off criteria) was used to analyze highly interconnected sub-clusters within the PPI network [24]. The DAVID web server was used to find the top enriched pathways by gene belonging to each cluster detected by MCODE. A network analyzer was used from Cytoscape software to detect the hub genes (nodes with maximum degree).

Total RNA-seq data retrieval and analyses

Total RNA sequencing data were retrieved from bio-project ID PRJNA882808; GEO: GSE213862. Genes with < 1 transcript per million (TPM) in $\geq 80\%$ of the samples were removed from the downstream analysis. The data were converted to zero median values with unit standard deviation (SD) to construct a matrix suitable for exploratory analyses. The differential gene expression between node negative and node positive samples was calculated using DESeq2. Data plotting and exploratory data analyses were performed using the “Tidyverse” and “ggpubr” packages in R (4.2.0).

MSR1 splice variant expression study

Transcripts for all splice variants of MSR1 were extracted from Ensembl database. TPM values were generated from RNA-Seq data and corresponding MSR1 transcripts were isolated. The box and whisker plot were successively generated using R ggplot and p -values were calculated using student’s t -test or \log_{10} transformed TPM values.

Immunohistochemistry (IHC) and multiplex-IHC staining and scoring

From each tumor tissue block, $3 \mu\text{m}$ -thick sections were prepared and dried in a $60 \text{ }^\circ\text{C}$ oven for 30 min. Single-marker immunohistochemistry (IHC) was performed in a Bond Max Automated Immunohistochemistry Vision Bio-system (Leica Microsystems GmbH, Wetzlar, Germany) according to standard protocols [25]. Markers specifically labeling myeloid cells and CAFs were myeloperoxidase (MPO) (polyclonal, RTU, DAKO) and smooth muscle actin- α (SMA- α) (clone 1A4, RTU, DAKO) respectively. Digital images of the stained slides were captured on an Aperio Versa 8 platform (Leica, Wetzlar Germany) at $20\times$ magnification and analyzed using QuPath software. Cells were scored at the invasive margin and tumor center [17]. The level of a cell type was determined based on the percentage of cells with at least the minimal level of marker expression intensity and thus considered “positive” by two independent experts.

In addition, multiplex IHC was performed to detect the co-localization of antigens using the following antibodies: Fibronectin-1 (clone HFN7.1, dilution 1:25, Medaysis); COL5A1 (clone E6U9W, dilution 1:50, Cell Signaling); FOXP3 (clone EP340, RTU, PathnSitu); CD25 (clone 4C9 RTU, Leica Biosystems); CD4 (clone EP204, RTU, Cell Signaling). Details of the methodology is described in supplementary Table S4.

Statistical analysis

All statistical analyses were performed in R (v4.2.1; R Foundation for Statistical Computing, Vienna, Austria).

Differential gene expression was determined by the LMM method with the assumption of several ROIs taken from the same patient tissue (random effect). *p* values were corrected for multiple tests using the false discovery rate method. Pearson's correlation of gene expression in each region of interest among different cell types were also measured in R and *p* value were tested, with confidence interval of 95%. The Kruskal–Wallis test was used to test for differences in gene expression across multiple ROI types, followed by the pair-wise Wilcoxon test. A *p*-value < 0.05 was considered statistically significant.

Results

Spatial transcriptome profiling of GB-OSCC tumor samples demarcates the tumor margin from the TME

To verify the data quality, two tests were performed to detect and discard outlier probes. Thus, 12 probes were eliminated as global outliers with low probe ratios. Using the remaining 8647 probes by spatial transcriptome profiling of 23 GB-OSCC tumor samples, 1812 genes were detected within the selected ROIs (LOQ threshold value = 2 to 227) with a detection rate < 1% to \geq 15% (Fig. S4A–C; Table S5 and S6). We observed that the ROIs in the tumor margins expressed the maximum number of genes at a distance less than 200 μ m between tumor and immune cells. Indeed, 20 of the tumor margin ROIs expressed > 15% of all the detected genes.

With the resulting data, we performed a Uniform Manifold Approximation and Projection (UMAP) based on the types of ROI: (a) invasive tumor margins with varying distance between tumor cells and immune cells (< 200 μ m, > 200, 300 and 500 μ m); (b) tumor center with a distance of 200 μ m between tumor and immune cells; (c) tumor cells only (without immune cells); (d) normal squamous epithelium. Spatial profiling across the tumor tissues revealed that although the tumor center and tumor margin ROIs had a scattered distribution throughout the dimensional reduction plot, they could be clearly distinguished based on PanCK+AOI and PanCK – AOI status (Fig. 1A). This segregation of ROIs based on PanCK+ and PanCK – status indicated differential spatial transcriptome profiles and showed that segregation depended on the type of segment rather than the lymph node status (Fig. 1B).

We also generated a heatmap based on 100 most highly variable genes (based on their CV% values) that formed two broad clusters of AOI: one covered tumor segments of the tumor margin (hereafter “tumor end”) and one covered TME segments (hereafter “stromal end”). These two clusters had a clear demarcation, representing a distinct gene expression profile of both regions (Fig. 1C).

Genes at tumor margins are crucial for nodal metastasis.

Having generated data on differential genes, we next queried the genes involved in lymph node metastasis. We performed a global differential gene expression analysis of the ROIs selected from lymph node-positive (*n* = 10) and lymph node-negative patients (*n* = 13). We detected 236 DEGs (Table S6, *p* < 0.05), of which 60 were significantly upregulated in the tumor margins of node-positive versus node-negative cases. Of these, the top 10 upregulated genes included: FN1, COL5A1, SFRP2, COL6A3, HLA-DRB5, SFRP2, LAMC2, PLAU, TPM1, and H3C2. The top 10 downregulated genes included: MSR1, SFRP1, HLA-DRB4, SERPINB2, GNLY, DEFB1, PRKACG, CEACAM3, COL4A6, and COL11A2 (Fig. 2A–D; Table S6).

We further analyzed the expression pattern of these DEGs in the ROIs of the tumor end and stromal end of the margins, while also considering the distance between the tumor and immune cells. Among the 60 upregulated in the node-positive cases, 17 genes (NDUFA2, P4HA1, NDUFA7, PTPN11, NDUFS8, TFDPI, GOT2, RELA, ADAM12, GPI, NPM1, MIF, ITGAV, TPM1, PLAU, COL5A1 and FN1) were at the tumor margins (< 200 μ m). Of these, FN1, COL5A1, TPM1, ADAM12, and P4HA1 were upregulated at the stromal end and the remaining genes (NDUFA2, NDUFA7, PTPN11, NDUFS8, TFDPI, GOT2, RELA, GPI, NPM1, MIF, ITGAV, PLAU) were upregulated at the tumor end of the invasive margins (Table S5). By contrast, we detected 106/176 that were downregulated at the tumor margins (< 200 μ m) (Table S5).

We next generated a volcano plot to visualize the top upregulated and downregulated genes between lymph node-positive and -negative ROIs (Fig. 2E). Overall, the nodal metastasis cases showed higher expression of genes including FN1, COL5A1, COL6A3, HLA-DRB5, SFRP2, ITGB6, ITGAV, and MIF at the invasive margins < 200 μ m, and the expression of the same genes gradually declined as the distance between tumor and immune cells increased. This pattern was not evident in the node-negative cases (Fig. 2E).

Having detected DEGs, we next performed a functional enrichment analysis using the DAVID service to understand the spatial impact of signaling pathways for cancer progression. When looking at the associations between gene expression and molecular functions, integrin binding and fibronectin binding (indicative of matrix remodeling) were enriched within the upregulated genes, while cytokine receptor activity (indicative of immune functions) was enriched among the downregulated genes. Some of the pathways associated with matrix remodeling found enriched were the ECM pathway, signal transduction pathway, and signaling by MET, while the

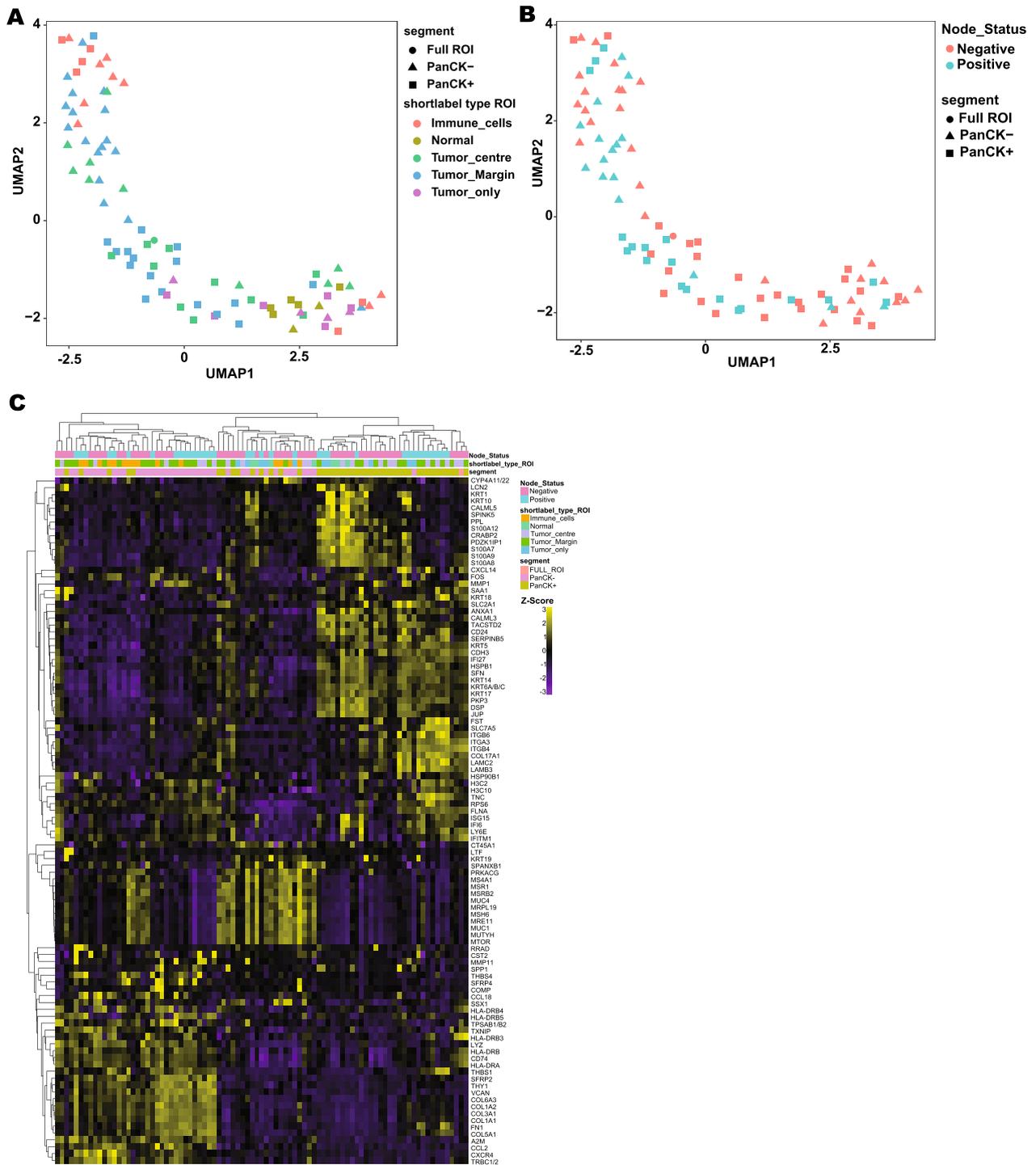


Fig. 1 Distance-wise spatial transcriptome. **A** Uniform Manifold Approximation and Projection (UMAP) dimensional reduction plot, generated using the Quartile 3 (Q3) method of normalized counts from the AOI profiled within the current spatial experiment. The UMAP generated based on the types of ROI spatially profiled across the tumor tissues where the shape stands for the type of the segment, i.e. either PanCK+ (tumor end) or PanCK- (stromal end) or Full ROI without any segmentation. **B** The UMAP generated based on the lymph node status, spatially profiled across the tumor tissues where the shape stands for the type of segment, i.e. either tumor end or stromal end or Full ROI without any segmentation. **C**, Hierarchical clustering of top 100 highly variable genes as calculated using the coefficient of variation (CV%); (N = 87 AOI from 23 patients). These 100 genes clustered the AOI into 2 broad groups. The 1st cluster mostly consisted of segments from the stromal end, and the 2nd cluster had segments from tumor end. The 2nd cluster was further sub-clustered into 2 groups, which was governed by the type of segments rather than the status of lymph nodes

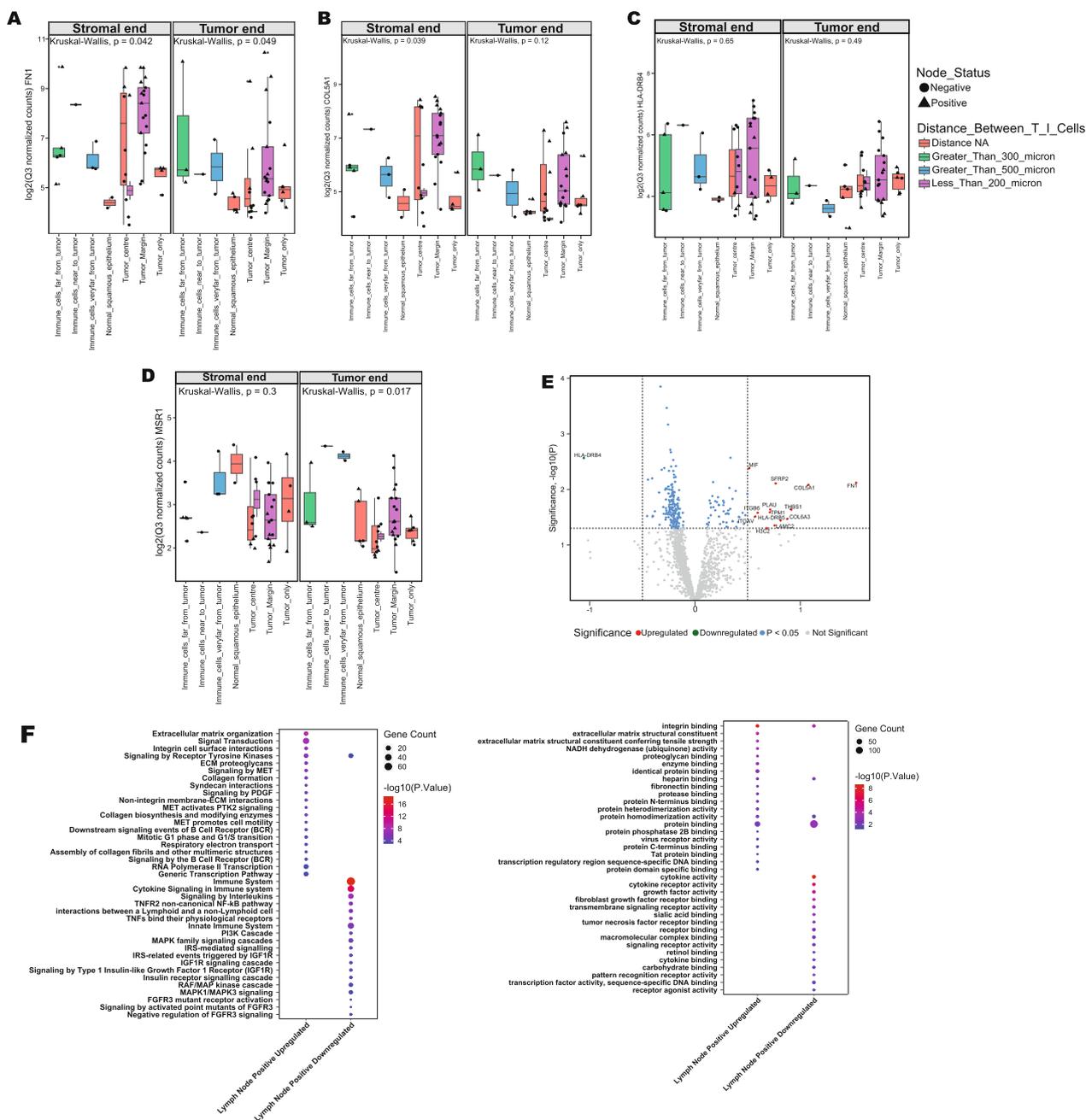


Fig. 2 Boxplot depicting the gene expression across multiple types of regions of interest spatially profiled for both lymph node-positive and negative patients. **A** FN1 gene was highly upregulated in lymph node-positive segments averaged across multiple types of ROI where increased expression at the tumor margin is observed with a significant difference with normal epithelial segments (pairwise Wilcoxon test p -value < 0.05). **B** Gene expression of COL5A1, second highly upregulated in lymph node-positive cases averaged across multiple types of ROI where a significant difference in expression was observed between stromal end, and cells in the TME at a distance $> 500 \mu\text{m}$. (pairwise Wilcoxon test p -value < 0.05) There was also differences between tumor margins and tumor regions (pairwise Wilcoxon test p -value < 0.01). **C** Gene expression of HLA-DRB4 was highly downregulated in lymph node-positive segments averaged across multiple types of ROI where no statistical significance was observed for expression across multiple ROI types in both tumor end and stromal end (Kruskal Wallis p value > 0.05). **D** Expression of MSR1, the second highly downregulated gene in lymph node-positive segments averaged across multiple types of ROI where statistically significant differences in its expression was observed as the distance between tumor and immune cells increased (pairwise Wilcoxon test p -value < 0.05), within stromal end segments. **E** Volcano plot for the differentially expressed genes between lymph node-positive ($n = 35$) against the lymph node-negative ($n = 52$) spatial regions of interest. **F** Top 20 DAVID functional enrichment terms for GO BP and REACTOME pathways respectively of significant DEGs (p -value < 0.05) between lymph node-positives vs negatives. GO gene ontology; DEG differentially expressed gene, DAVID Database for Annotation, Visualization, and Integrated Discovery, BP biological processes

downregulated pathways included signaling by interleukins and the TNFR2 non-canonical NF- κ B pathway, amongst others (Fig. 2F). Taken together, there was disruption in matrix biology and immune function in cases with nodal metastasis.

ECM genes distributed to the tumor margins assist in tumor metastasis.

To understand the mechanism of nodal metastasis in GB-OSCC, we first validated that the tumor margins are the site of metastatic signature. To do so, we evaluated the spatial distribution of known metastasis markers, including PDPN, ZEB1, and LAMB3. PDPN and LAMB3 were over-expressed at the tumor end while

ZEB1 was more expressed at the stromal end of the tumor margins (Fig. 3A–C).

We then performed a spatial study into the most significantly upregulated (FN1 and COL5A1) and down-regulated genes (MSR1 and HLA-DRB4) in tumor specimens taken from patients with nodal metastasis. We observed that FN1 and COL5A1 – which encode Fibronectin 1 and Collagen Type V Alpha 1 Chain respectively—were very highly expressed at a distance <200 μ m and at the stromal end (Fig. 2A, B; Table S5). Protein–protein interaction analysis revealed that although FN1 and COL5A1 belong to different functional clusters, inter communication between both

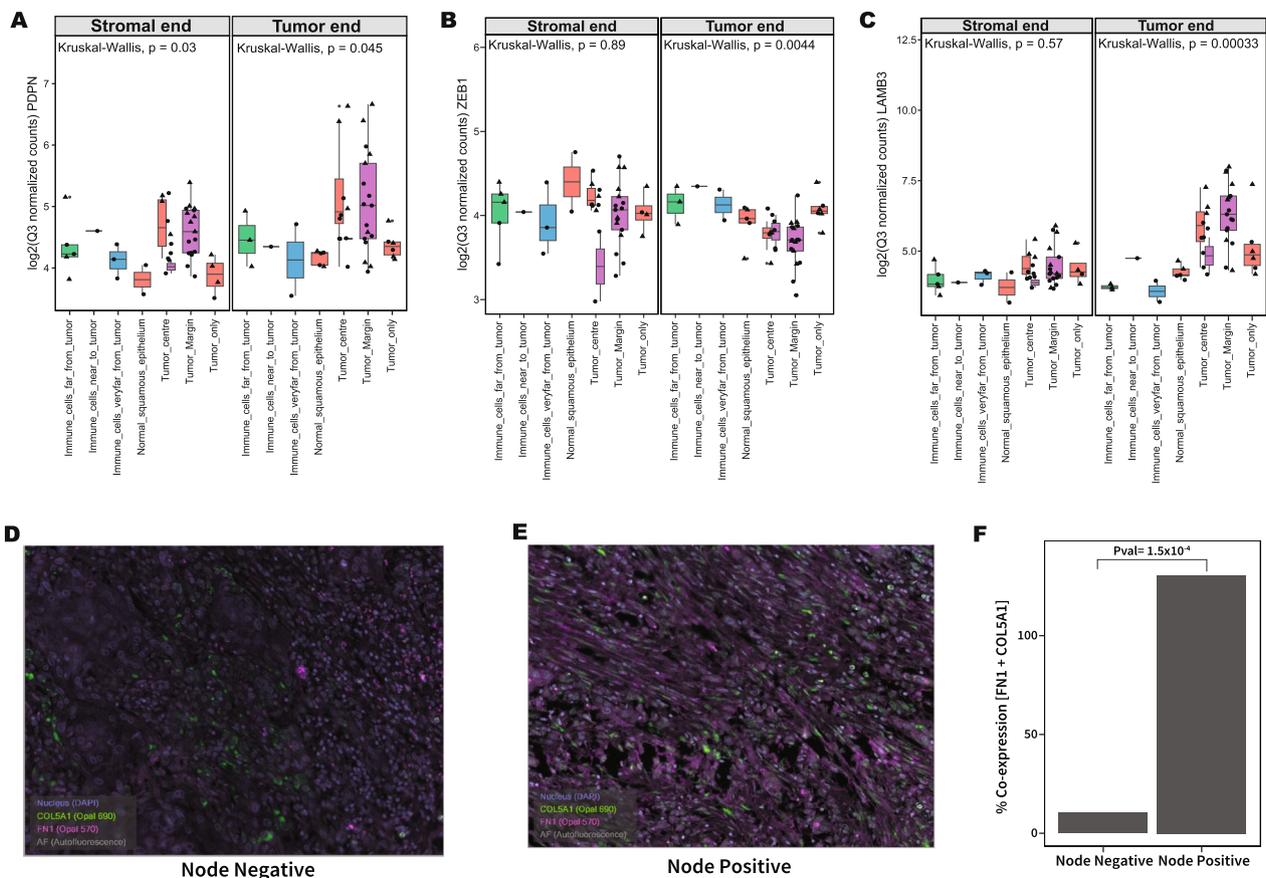


Fig. 3 Boxplot depicting the metastatic gene expression across multiple types of ROI spatially profiled for both lymph node-positive and negative patients. **A** Expression of LAMB3 gene averaged across multiple types of ROI where high expression was observed at the tumor end, along with significant differences with normal epithelial segments (pairwise Wilcoxon test p-value < 0.001). **B** Gene expression of PDPN averaged across multiple types of ROI where high expression was observed at the tumor end, along with significant differences in expression between tumor margin and other ROI types like tumor and tumor centre (pairwise Wilcoxon test p-value < 0.05). **C** Gene expression of ZEB1 averaged across multiple types of ROI where high expression was observed at the stromal end, where no statistical significance was observed for differences in expression across multiple ROI types (Kruskal Wallis p-value > 0.05). However, the segment of tumor end was supported by high statistical significance for the differences in gene expression between the different ROI types (pairwise Wilcoxon test p-value < 0.01). Co-localization study by multiplex IHC. Co-expression of FN1 and COL5A1 in **D** node-negative; **E** node-positive cases. **F** Nodal metastasis showed higher FN1 and COL5A1 co-expression. n = 8, p-value = 1.5×10^{-4}

clusters signified a functional correlation between them for the organization of the ECM (Fig.S5).

Multiplex IHC revealed co-localization of FN1 and COL5A1 in the same fibroblast-like cells of the TME at the invasive margins. The number of cells showing co-existence of the two proteins was much higher in cases with nodal metastasis (Fig. 3D–F).

Low MSR1 and HLA-DRB4 expression—which encode Macrophage Scavenger Receptor 1 and Major Histocompatibility Complex, Class II, DR Beta 4 respectively—was observed at margins, at a distance < 200 μm (Fig. 2C, D). Of note, bulk RNA-sequencing data from a cohort of GB-OSCC (GSE213862) belonging to our hospital patient samples showed upregulation of MSR1. Bulk RNA-Seq data also did not identify FN1 or COL5A1 in the top upregulated genes in cases with nodal metastasis and HLA-DRB4 was not in the list of DGEs (Table S7, Fig. S6A). Such downregulation of the MSR1 gene occurs due to the isoform-specific dominant negative effect of MSR1 (Fig. S6B). Taken together, FN1 and COL5A1 (like other ECM genes) at the tumor margins assist in nodal metastasis. Therefore, it is evident that spatial transcriptomics has the potential to be more precise and refine our understating of disease progression.

GSVA analysis of HALLMARK gene sets

Our next set of analyses aimed to understand the impact of the spatially distributed gene signatures on biological pathway regulation. To do so, we performed a GSVA analysis, for the HALLMARK gene sets for the different ROI types based on the distance between tumor and immune cells status to detect signaling pathway enrichment within specific segments. We observed a clear difference in signaling pathway regulation in both node-positive and node-negative patients, based on tumor and stromal segmentation. Specifically, we saw that DNA repair, unfolded protein response, mTORC1, E2F, and oxidative phosphorylation pathways were more active in the tumor end of the tumor margins in lymph node-positive patients. Angiogenesis, protein secretion, apical surface, and NOTCH signaling pathways were instead more

active in the stromal end of the tumor margins (Fig. 4A, B).

GSVA analysis of OSCC-specific functional gene sets

Next, we wanted to determine some of the relevant OSCC-specific functional genesets related to tumor progression and metastasis in all ROI types from the four broad AOI groups. Here, the signaling pathways differed according to the distance between tumor and immune cells (Fig. 4C, D). The Partial EMT geneset (p-EMT) contains genes associated with ECM components such as matrix MMPs, laminins, and integrins and that regulate tumor invasion and metastasis [20]. We observed that the p-EMT geneset was overexpressed at the margin at a distance < 200 μm for lymph node-positive cases (Fig. 4C). Importantly, we saw that CAFs were predominantly functional at the stromal end (Table S8). We thus conclude that tumor margins are the active site for nodal metastasis. Also, CAFs at the stromal end could assist the ECM formation.

Cancer-immune cell type abundance in the OSCC ROIs using Spatial Decon

Using the Spatial Decon R package, we were able to deconvolute the abundance of 14 different immune cell types by leveraging the safeTME matrix provided by the Cell Profile Library database (<https://github.com/Nanostring-Biostats/CellProfileLibrary>) (Table S8). We observed that cell types like neutrophils were more prevalent at the tumor end while CD4-T cells, endothelial cells, fibroblasts, NK cells, and Treg cells were prevalent at the stromal end (Fig. 5A). Cases with lymph node metastasis comprised more Treg cells and fibroblasts at the stromal end while CD8+ and CD4+ memory cells were more abundant at an extended distance (Fig. 5B).

Our multiplex IHC data further confirmed the increased presence of Treg cells (CD4+CD25+FOXP3) in the tumor margins in cases having node metastasis (Fig. 5C–E).

Using the OSCC single-cell RNA-Seq data (GSE103322) already available [21], we were also able to deconvolute certain cell types specific to OSCC which

(See figure on next page.)

Fig. 4 GSVA enrichment analysis in Hallmarks and functional genesets specific to OSCC. Heatmap plot representing the value of GSVA score calculated for each type of ROI, belonging to 4 groups of patients. **A** Lymph node-positive–tumor end, lymph node positive–stromal end, **B** Lymph node negative–tumor end and lymph node-negative- stromal end). Gradient-wise change in the regulation status of the Hallmark genesets is observed as the distance between tumor and immune cells increases. Positive GSVA scores represent genes in the HALLMARK gene set that are positively enriched while negative GSVA scores represent genes in the HALLMARK gene set that are negatively enriched. Gradient-wise change in the regulation status of the OSCC-specific expression programs can be observed as the distance between tumor and immune cells increases. **C** Lymph node-positive–tumor end, lymph node positive–stromal end, **D** Lymph node negative– tumor end, and lymph node-negative- stromal end. Positive GSVA scores represent genes in the expression program that are positively enriched while negative GSVA scores represent genes in the expression program that are negatively enriched

are not available in the safeTME cell profile. Here, we saw that intermediary resting fibroblasts and myocytes were prevalent in the tumor end segments while CD4+, CD8+, and CD8+ exhausted T cells were more abundant at the stromal end (Fig. 6A, B).

H&E and IHC studies for cell type abundance at tumor margins

We reviewed the H&E images to evaluate the abundance of cell types of our interest. We saw that neutrophils, CAFs, and collagens were present at the tumor margins (Fig. 7A). We confirmed this finding by IHC, showing that the tumor margins had a high density of MPO (a marker of myeloid cells) and SMA α (a marker of CAFs) protein expression in node-positive patients (Fig. 7B–E).

Pearson's correlation studies of FN1 and COL5A1 with fibroblasts and neutrophils

To understand how our genes of interest correlated with the regionally abundant fibroblast and neutrophils, Pearson's correlation studies was performed. We observed both FN1 ($R=0.87$, $p<2.2e-16$) and COL5A1 ($R=0.92$, $p<2.2e-16$) had a positive correlation with the stromal fibroblasts. However, there was no correlation of neutrophils either with FN1 ($R=0.014$, $p<0.95$) or COL5A1 ($R=0.017$, $p<0.93$) (Fig. 8). We also found no correlation between HLA-DRB4 and other immune cell types (Table S9).

Discussion

The 1-mm area of the tumor-stromal interface, called the tumor margin, supposedly decides the fate of tumor progression [17, 26]. In this study, we annotated the GB-OSCC tumor margin at the genomic level, in order to identify prognostic biomarkers and possible targets. We uncovered the differential spatial transcriptomic profiles of >1800 genes at the tumor margin at variable distances of <200 μm to >500 μm between tumor and immune cells. We also captured the transcriptomic profiles at the tumor centers and in the tumor cells and normal epithelium, which revealed differential gene distribution within the different regions. Due to the enrichment of gene signatures at the invasive edge, we focused our explorations on the differences in gene expression between tumor and

immune cells separated by $\leq 200 \mu\text{m}$, with a focus on the tumor (AOI-1) and stromal ends (AOI-2). In this way, we could understand how the cells in these discrete regions differ in terms of their role in tumorigenesis and lymph node metastasis.

Of the 23 cases of GB-OSCC included in this study, 10 had cervical lymph node metastasis while 13 were node-negative based on pathological staging on the resected specimens. We annotated the differential gene expression pattern according to the nodal status, and found 236 DEGs between node-positive and node-negative patients. Of these, only 60 genes were upregulated in node-positive cases while the rest were downregulated. Among these 60 genes, top ranking DEGs included FN1 ($p=0.007$) and COL5A1 ($p=0.008$), which have known roles in regulating the ECM. However, FN1 and COL5A1 were not the top upregulated genes in a previous bulk RNA-Seq study (GSE213862). While this contradictory finding requires validation, it should be noted that true signals can be obscured by an average gene expression profile from bulk RNA-Seq while spatial transcriptomics offers precision and accuracy.

Other notable findings included high LAMB3 and PDPN expression at the tumor end, and high ZEB1 expression at the stromal end of the tumor margin in cases with metastasis. These findings suggest that the tumor margin is an active region for metastasis and that both segments of the tumor margin have roles in tumor progression. Indeed, the invasive edge is known for its metastatic function, and LAMB3, PDPN, and ZEB1 activity have all been implicated in this region in other cancers [27–29]. GSVA enrichment analysis for functional gene sets specific to OSCC also illustrated the importance of the tumor margin, where a gradient-wise change in the regulation status of the OSCC-specific expression was observed for all sample types, and there were observable differences in pathway impacts even within the tumor margins.

We consider that the spatial and functional organization of ECM molecules like FN1 and collagens like COL5A1 are likely key targets in oral cancer progression. FN1 was highly expressed at the tumor margin (particularly the stromal end) where the distance between the immune cells and tumor cells was <200 μm . We infer that this proximity of FN1 to the stroma allows direct

(See figure on next page.)

Fig. 5 Immune cell deconvolution in tumor end and stromal end segments of OSCC. **A** Proportions of 14 different cell type population in the selected ROI across 23 OSCC patients. **B** Immune cell populations in the OSCC cohort were visualized for the differences in abundance as the distance between tumor and immune cells increased. Bubble plot showing an abundance of immune cell types in each AOI calculated by spatial deconvolution. Identification of Treg cells by signatory protein markers. CD4, CD25 and FOXP3 triple positive status in **C** node negative; **D** node positive cases. **E** Nodal metastasis showed higher triple positive (CD4, CD25 and FOXP3) status. $n=8$, $p\text{-value}=9.3 \times 10^{-4}$

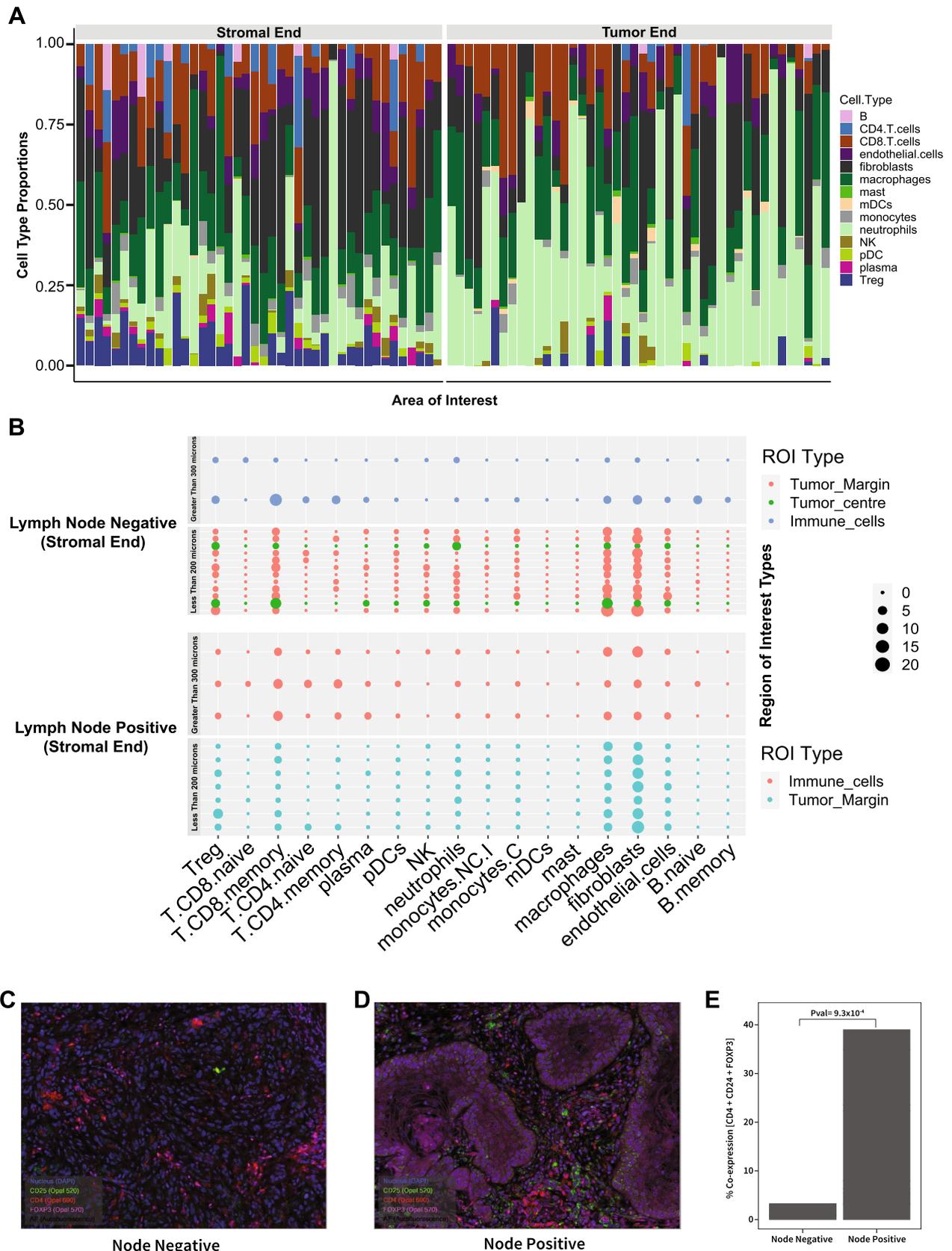


Fig. 5 (See legend on previous page.)

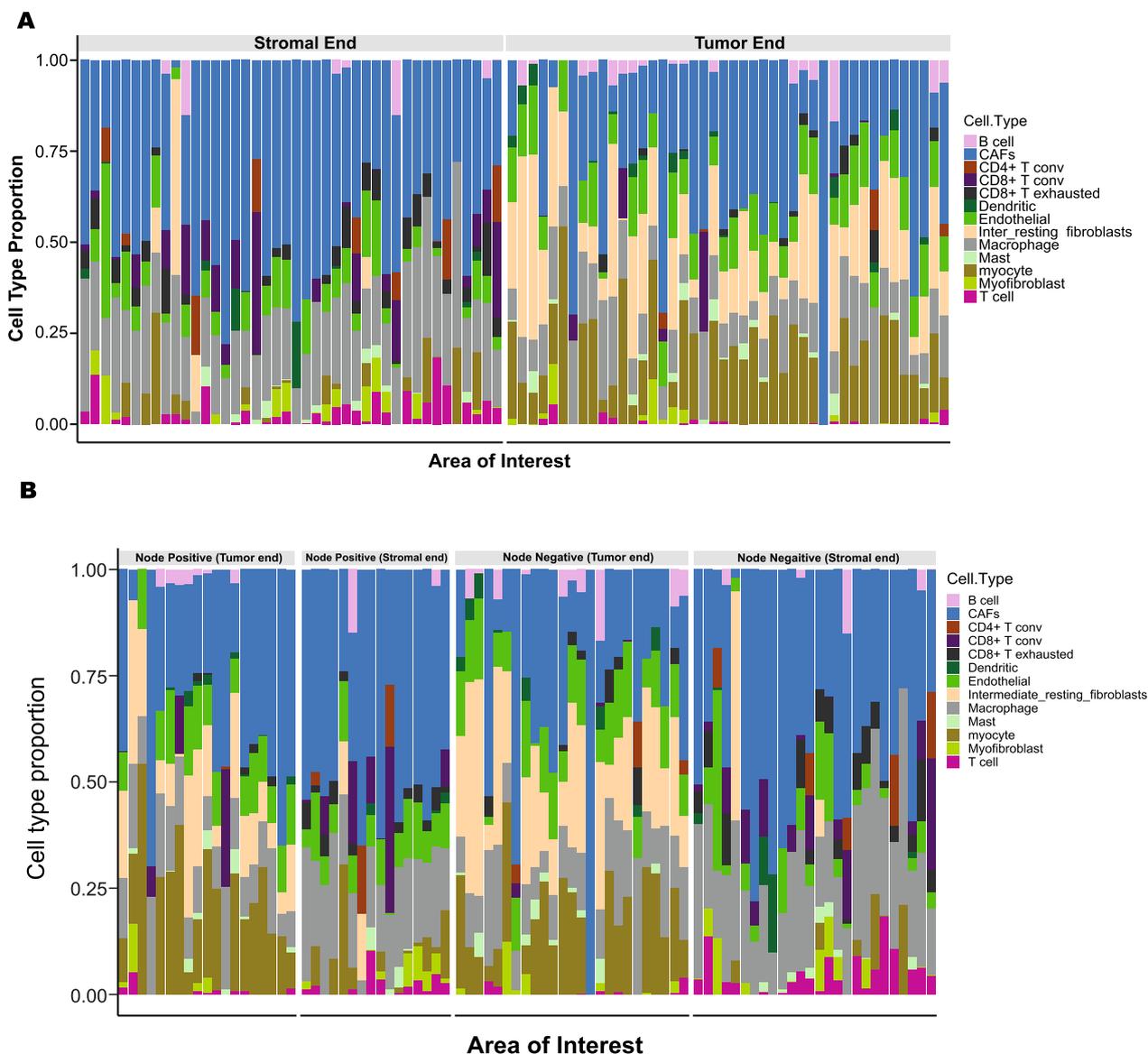


Fig. 6 OSCC-specific cell type deconvolution. OSCC-specific cellular composition within each spatially profiled ROI were deconvoluted using a signature matrix, generated from a single-cell RNAseq dataset which included OSCC cells (GSE103322). **A** Proportions of 13 different OSCC-specific cell type population in the segmented spatially profiled at tumor end and stromal end across 23 OSCC patients. **B** Immune cell deconvolution in stromal end and tumor end of node positive and node negative patients. Intermediate resting fibroblasts are found more in abundance within lymph node-positive patients

remodeling of the ECM in the TME, where FN1 can bind to cell surfaces and various cell components in the ECM. Interestingly, FN1 and integrins have been proposed as potential biomarkers for tongue/floor-of-mouth cancers [30]. Our study also showed metastasis to lymph nodes is driven by FN1 activation, at least by three upregulated integrin gene signatures, namely ITGA5, ITGB6, and ITGAV. Our functional enrichment analysis also showed that integrin-mediated signaling pathways had an impact on nodal metastasis in GB-OSCC. Collagen

subtypes COL5A1, COL5A2, COL6A3, and COL27A1 were upregulated in patients with nodal metastasis, while COL11A2 and COL4A6 were downregulated. In particular, COL5A1 may have suggestive significance in other cancers [31, 32]. In our study we noticed COL5A1 to be highly expressed in the same region as FN1, and that both exhibited a significant, positive correlation with fibroblasts in the TME at the stromal end. Since FN1, and COL5A1 were located at the stromal end, we hypothesize that the two jointly work in tandem to remodel the ECM.

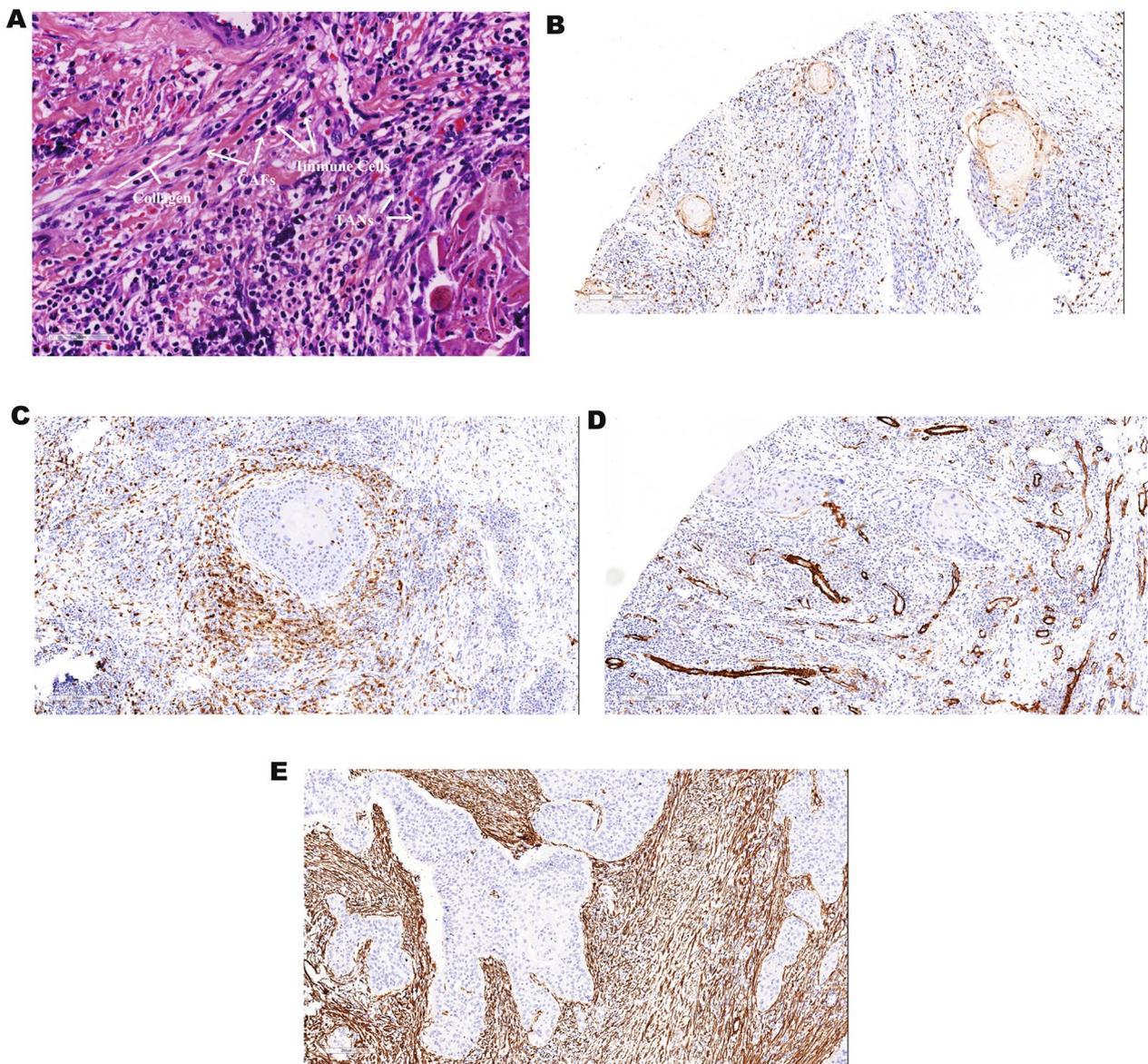


Fig. 7 Determination of TANs and CAFs at tumor margin. **A** H&E images for identification of CAF, TAN, Collagen, and immune cell morphologies in the tumor margins. **B, C** Comparative study by IHC (Myeloperoxidase (MPO)) for TANs in **B** node-negative cases; **C** node-positive cases. **D, E** Comparative study by IHC (Smooth muscle antibodies (SMA α)) for CAFs in **D** node-negative cases; **E** node-positive cases

In support of this, we have performed multiplex IHC using antibodies to FN1 and COL5A1 and shown that the two are co-localized in the same cells. We propose, therefore, that FN1 and COL5A1 collaborate to actively remodel the matrix for tumor metastasis. Meanwhile, a correlation between CAFs and FN1 at the tumor margins was independent of nodal metastasis, possibly implying that CAFs can produce FN1 upon activation and following specific signals, in node-positive cases. Indeed, we observed the presence of CAFs on histology sections at the tumor margins. We consider that the co-localization

of CAFs with FN1 and COL5A1 signifies that CAFs are proficient in modulating the ECM through the activation of FN1 and COL5A1.

Previous studies have also shown that integrin-mediated signaling enhances CAFs and M2 macrophage participation in tumor growth [33]. In the TME, CAFs interact with immune cells and other components via the secretion of various cytokines, chemokines, growth factors, and exosomes, to produce an immunosuppressive TME that enables cancer cells to evade the immune surveillance [33, 34]. Our findings are in concurrence with

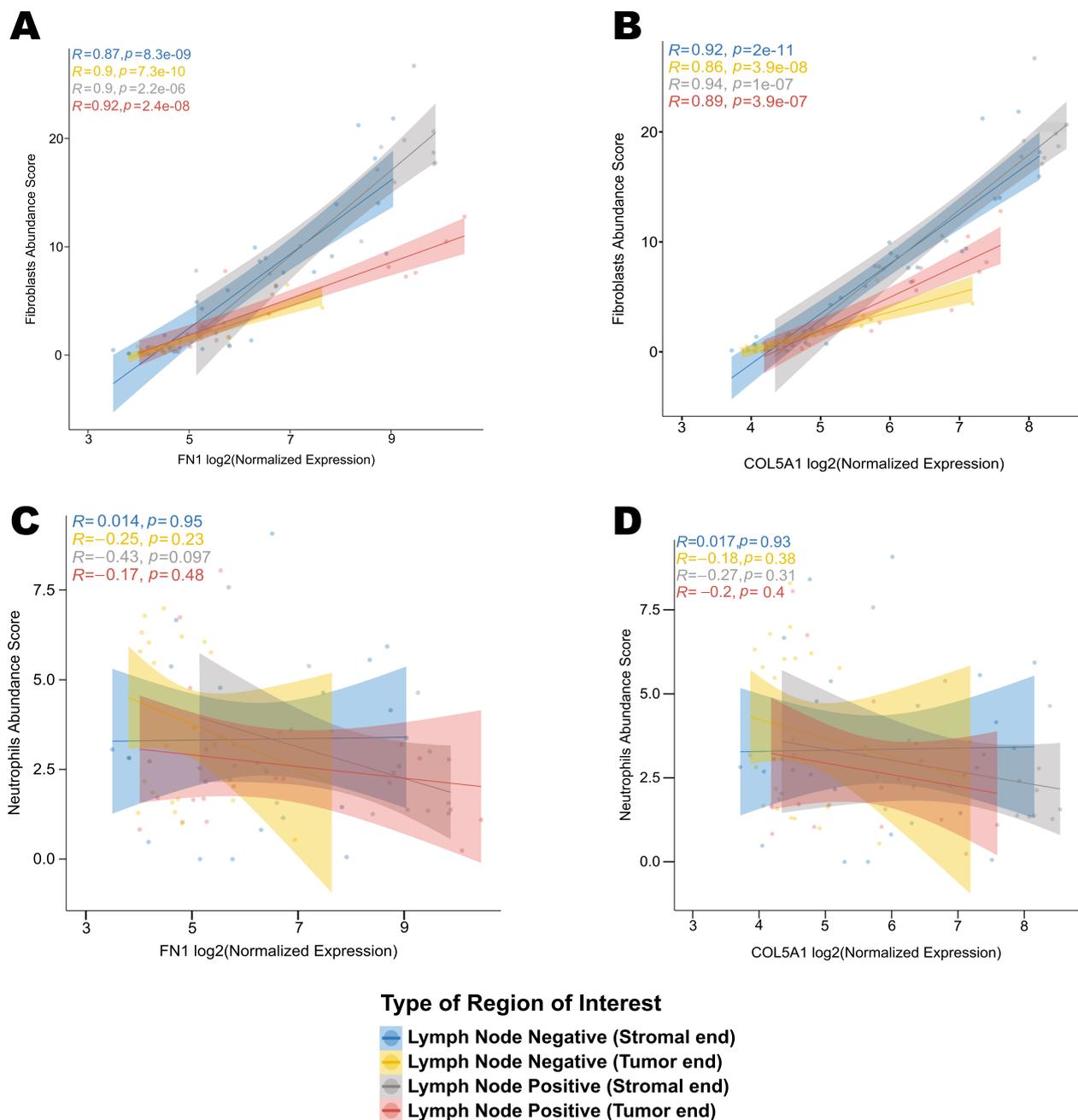


Fig. 8 Correlation of ECM components FN1 and COL5A1 with fibroblast and neutrophils. **A, B** Positive correlation of FN1 ($R=0.87, p < 2.2e-16$) and COL5A1 ($R=0.92, p < 2.2e-16$) respectively with fibroblasts in the TME. **C, D** However, there was no correlation with neutrophils

an immune-suppressed invasive margin, particularly in node-positive cases, as evidenced by the enrichment of Treg cells at the stromal end. To confirm this, multiplex IHC for Tregs was done using antibodies to CD4, CD25 and FOXP3. Cases with node metastasis showed higher numbers of Tregs—that is cells showing co-localization of all the 3 markers.

Meanwhile, cells enriched at the tumor ends of the invasive margins included intermediate fibroblasts, myocytes, and neutrophils. Intermediate resting fibroblasts are likely CAF precursors [35]. Moreover, recent findings showed that CAFs can also originate from smooth muscle and pericytes [36]. With this in mind, we posit that there is a sequential conversion of myocytes and/or intermediate resting fibroblasts

to form CAFs at the tumor end that migrate to the adjacent stromal end for FN1-mediated ECM modulation. In support of this, we noticed that the population of intermediate resting fibroblasts was higher at the tumor end in node-negative cases than in node-positive cases, signifying that the intermediate resting fibroblasts converted to mature CAFs with disease progression.

Neutrophils were abundant at the tumor end, though there was no correlation between neutrophils and FN1 and COL5A1 indicating that if neutrophils assist in ECM formation, they do so independently. Neutrophils have inhibitory effects on T-cells, promoting metastasis in other cancers [37]. Consistently, we found a higher T-cell population together with exhausted T-cells at the stromal end, suggesting that neutrophils at the tumor end might retard immune infiltration and promote tumor invasion. A weakened immune system facilitates tumor invasion and metastasis in all cancer types. We found HLA-DRB4 and MSR1 encoding genes were significantly downregulated at the stromal end in cases with nodal metastasis. In general, Class II HLA binds peptides derived from tumor cells and are presented on the cell surface for recognition by CD4 T-cells. A previous study reported that the expression of HLA-DRB4, a Class II HLA, is directly proportional to patient survival [38]. Although we saw that CD4 T-cells were abundant at the tumor margins likely due to HLA-DRB4 downregulation, we saw no correlation between HLA-DRB4 and other immune cell types except monocytes in node-negative cases. Other class II HLAs like HLA-DPB1 and HLA-DQA2 were also downregulated in node-positive cases while HLA-DRB5 was upregulated. The decreased expression of HLA-DRB4 at the tumor margins in node-positive cases might facilitate the immune escape mechanism and promote matrix formation by FN1, COL5A1 and related genes for tumor progression.

MSR1 has been implicated in promoting tumor metastasis in various cancers [39]. Data from our total RNA-seq analysis concur with this finding, with significantly higher MSR1 expression in node metastasized tumor samples. Yet by contrast, data from our spatial transcriptomic analysis showed downregulation of the MSR1 gene at the tumor margins of lymph node-positive cases compared to lymph node-negative cases. On reviewing the possible causes of downregulation in the lymph node-positives in spatial context, we observed isoform-specific dominant negative effect of MSR1. The MSR1 gene encodes three different isoforms created by alternative splicing. The third isoform has a dominant negative effect when co-expressed with other isoforms. In our study, isoform 3 was over-expressed

in node-positive cases, which we would expect to suppress isoforms 1 and 2 in the spatial context. By contrast, one study observed that MSR1 and TLR7 could be anti-metastatic [39], which concurs with our findings.

Conclusion

We found that a 200 μm tumor-stromal interface was the major site of biological activity in the GB-OSCC invasive margin. Immune and stromal cells at the tumor end of this interface, almost touching the tumor cells, exhibited a distinct gene expression pattern compared to that of cells situated at the stromal end. At the stromal end, fibronectin (FN1) and collagen-related gene expression modulated the extracellular matrix (ECM), while immune suppression fostered lymph node metastasis. At the tumor end, intermediate fibroblasts and myocytes served as precursor cancer-associated fibroblasts (CAFs) that migrated to the stromal end and contributed to FN1-mediated ECM modulation. Neutrophils at the tumor end were also associated with immune suppression and lymph-node metastasis.

CAFs have attracted much attention in recent years as one of the main components of the TME. They interact with cancer and immune cells, participate in ECM remodeling, and help mediate many immune responses. Given their importance, the next step is now to characterize CAFs in oral cancer and their interactions with immune cells so that we can predict immune efficacy and provide tailored combination immunotherapy. This insight might eventually lead to targeting CAFs as a potential clinical application.

Unfortunately, there are currently no existing experimental models that mimic the invasive margins of tumors. Our aim is to continue to advance our understanding of the spatial location and functional organization of the cell types at the invasive margins but a key issue is that as repeat biopsies are not possible from the same site with disease progression, the conversion between cell types cannot be tracked. A strategy for the future with great potential is to use organoid-in-chip models. Nevertheless, our initial data serve as a strong and impactful resource for future research into the cross talk between the GB-OSCC tumor and its TME.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-024-05511-1>.

- Additional file 1.
- Additional file 2.
- Additional file 3.
- Additional file 4.
- Additional file 5.

Additional file 6.
 Additional file 7.
 Additional file 8.
 Additional file 9.
 Additional file 10.
 Additional file 11.
 Additional file 12.
 Additional file 13.
 Additional file 14.
 Additional file 15.

Acknowledgements

The authors would like to acknowledge the Department of Biotechnology, Government of India, New Delhi, for partial funding under the Systems Medicine Cluster project. The authors also thank Mr. Samrat Roy, Mr. Tapash Kumar Giri Sarkar and Mr. Gopal Chakraborty, Tata medical center for their technical support.

Author contributions

SS, GM, SD, HD and MM designed the experiments. MM, VB, SS and GM performed the experiments. MM, VB, SS, GM and SB analyzed the data. SD, DB, and HD provided access to clinical samples. HD, GM and DKM provided intellectual discussion. MM, SS, GM wrote the manuscript. All authors critically revised and approved the manuscript.

Funding

This study was partially supported by the Department of Biotechnology, Government of India, New Delhi (Project Reference: No./BT/Med-II/NIBMG/SyMeC/2014/Vol II).

Availability of data and materials

Data information has been mentioned in the corresponding section of the manuscript. However, if additional information is required, it will be shared upon reasonable request made to the corresponding author.

Declarations

Ethics approval and consent to participate

This project was approved by the Tata Medical Center Institutional Review Board (IRB) (Ref. no. EC/GOVT/23/17).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Tata Medical Center, 14 MAR (E-W), New Town, Rajarhat, Kolkata, WB 700160, India. ²Tata Consultancy Services (TCS), Kolkata, WB, India. ³Medica Superspecialty Hospital, 127, Eastern Metropolitan Bypass, Nitai Nagar, Mukundapur, Kolkata, WB 700099, India. ⁴theraCUES Innovations Pvt Ltd., Bangalore, Karnataka 560092, India. ⁵Tata Translational Cancer Research Centre (TTCRC), 14 MAR (E-W), New Town, Rajarhat, Kolkata, WB 700160, India. ⁶Narayana Superspecialty Hospital, 120, 1, Andul Rd, Shibpur, Howrah, WB 711103, India.

Received: 4 March 2024 Accepted: 15 July 2024

Published online: 13 August 2024

References

- Mathur P, Mehrotra R, Fitzmaurice C, et al. Cancer trends and burden in India—Authors' response. *Lancet Oncol*. 2018;19(12): e664. [https://doi.org/10.1016/S1470-2045\(18\)30857-X](https://doi.org/10.1016/S1470-2045(18)30857-X).
- Mandlik DS, et al. Squamous cell carcinoma of gingivobuccal complex: literature, evidences and practice. *J Head Neck Phys Surg*. 2018;6(1):18–28. https://doi.org/10.4103/jhnps;jhnps_19_18.
- Singhania V, et al. Carcinoma of buccal mucosa: a site specific clinical audit. *Ind J Can*. 2015;52:605–10. <https://doi.org/10.4103/0019-509X.178383>.
- Manjula BV, Augustine S, Selvam S, Mohan AM. Prognostic and predictive factors in gingivo buccal complex squamous cell carcinoma: role of tumor budding and pattern of invasion. *Indian J Otolaryngol Head Neck Surg*. 2015;67(Suppl 1):98–104. <https://doi.org/10.1007/s12070-014-0787-2>.
- Baghban R, Roshangar L, Jahanban-Esfahlan R, et al. Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Commun Signal*. 2020;18(1):59. <https://doi.org/10.1186/s12964-020-0530-4>.
- Gou Z, Li J, Liu J, Yang N. The hidden messengers: cancer associated fibroblasts-derived exosomal miRNAs as key regulators of cancer malignancy. *Front Cell Dev Biol*. 2024;12:1378302. <https://doi.org/10.3389/fcell.2024.1378302>.
- Tagirasa R, Yoo E. Role of serine proteases at the tumor-stroma interface. *Front Immunol*. 2022;13: 832418. <https://doi.org/10.3389/fimmu.2022.832418>.
- Pruitt HC, Guan Y, Liu H, et al. Collagen VI deposition mediates stromal T cell trapping through inhibition of T cell motility in the prostate tumor microenvironment. *Matrix Biol*. 2023;121:90–104. <https://doi.org/10.1016/j.matbio.2023.06.002>.
- Efthymiou G, Saint A, Ruff M, Rekad Z, Ciais D, Van Obberghen-Schilling E. Shaping up the tumor microenvironment with cellular fibronectin. *Front Oncol*. 2020;10:641. <https://doi.org/10.3389/fonc.2020.00641>.
- Kim SW, Kim CW, Moon YA, Kim HS. Reprogramming of tumor-associated macrophages by metabolites generated from tumor microenvironment. *Anim Cells Syst (Seoul)*. 2024;28(1):123–36. <https://doi.org/10.1080/19768354.2024.2336249>.
- Belhabib I, Zaghoudi S, Lac C, Bousquet C, Jean C. Extracellular matrices and cancer-associated fibroblasts: targets for cancer diagnosis and therapy? *Cancers (Basel)*. 2021;13(14):3466. <https://doi.org/10.3390/cancers13143466>.
- Peng Z, Tong Z, Ren Z, Ye M, Hu K. Cancer-associated fibroblasts and its derived exosomes: a new perspective for reshaping the tumor microenvironment. *Mol Med*. 2023;29(1):66. <https://doi.org/10.1186/s10020-023-00665-y>.
- Lei X, Lei Y, Li JK, et al. Immune cells within the tumor microenvironment: biological functions and roles in cancer immunotherapy. *Cancer Lett*. 2020;470:126–33. <https://doi.org/10.1016/j.canlet.2019.11.009>.
- Chaudhary A, Bag S, Arora N, Radhakrishnan VS, Mishra D, Mukherjee G. Hypoxic transformation of immune cell metabolism within the microenvironment of oral cancers. *Front Oral Health*. 2020;1: 585710. <https://doi.org/10.3389/froh.2020.585710>.
- Arora R, Cao C, Kumar M, et al. Spatial transcriptomics reveals distinct and conserved tumor core and edge architectures that predict survival and targeted therapy response. *Nat Commun*. 2023;14(1):5029. <https://doi.org/10.1038/s41467-023-40271-4>.
- Hutchins GG, Grabsch HI. Molecular pathology—the future? *Surgeon*. 2009;7(6):366–77. [https://doi.org/10.1016/s1479-666x\(09\)80112-1](https://doi.org/10.1016/s1479-666x(09)80112-1).
- Hendry S, Salgado R, Gevaert T, et al. Assessing tumor-infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the international immuno-oncology biomarkers working group: part 2: TILs in melanoma, gastrointestinal tract carcinomas, non-small cell lung carcinoma and mesothelioma, endometrial and ovarian carcinomas, squamous cell carcinoma of the head and neck, genitourinary carcinomas, and primary brain tumors. *Adv Anat Pathol*. 2017;24(6):311–35. <https://doi.org/10.1097/PAP.0000000000000161>.
- Ortogerio N et al. GeomxTools: NanoString GeoMx Tools. R package version 3.4.0. 2023.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. <https://doi.org/10.1186/s13059-014-0550-8>.

20. Griswold M et al. SpatialDecon: Deconvolution of mixed cells from spatial and/or bulk gene expression data. R package version 1.10.0. 2023.
21. Puram SV, Tirosh I, Parkih AS, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell*. 2017;171(7):1611-1624.e24. <https://doi.org/10.1016/j.cell.2017.10.044>.
22. Hännelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7. <https://doi.org/10.1186/1471-2105-14-7>.
23. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–25. <https://doi.org/10.1016/j.cels.2015.12.004>.
24. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*. 2003;4:2. <https://doi.org/10.1186/1471-2105-4-2>.
25. Mukherjee G, Bag S, Chakraborty P, et al. Density of CD3+ and CD8+ cells in gingivo-buccal oral squamous cell carcinoma is associated with lymph node metastases and survival. *PLoS ONE*. 2020;15(11):e0242058. <https://doi.org/10.1371/journal.pone.0242058>.
26. Phanthunane C, Wijers R, de Herdt M, et al. B-cell clusters at the invasive margin associate with longer survival in early-stage oral-tongue cancer patients. *Oncoimmunology*. 2021;10(1):1882743. <https://doi.org/10.1080/2162402X.2021.1882743>.
27. Liu L, Jung SN, Oh C, et al. LAMB3 is associated with disease progression and cisplatin cytotoxic sensitivity in head and neck squamous cell carcinoma. *Eur J Surg Oncol*. 2019;45(3):359–65. <https://doi.org/10.1016/j.ejso.2018.10.543>.
28. Wang X, Li W, Bi J, et al. Association of high PDPN expression with pulmonary metastasis of osteosarcoma and patient prognosis. *Oncol Lett*. 2019;18(6):6323–30. <https://doi.org/10.3892/ol.2019.11053>.
29. Horny K, Sproll C, Peiffer L, et al. Mesenchymal-epithelial transition in lymph node metastases of oral squamous cell carcinoma is accompanied by ZEB1 expression. *J Transl Med*. 2023;21(1):267. <https://doi.org/10.1186/s12967-023-04102-w>.
30. Yen CY, Huang CY, Hou MF, et al. Evaluating the performance of fibronectin 1 (FN1), integrin $\alpha 4 \beta 1$ (ITGA4), syndecan-2 (SDC2), and glycoprotein CD44 as the potential biomarkers of oral squamous cell carcinoma (OSCC). *Biomarkers*. 2013;18(1):63–72. <https://doi.org/10.3109/1354750X.2012.737025>.
31. Boguslawska J, Kedzierska H, Poplawski P, Rybicka B, Tanski Z, Piekielko-Witkowska A. Expression of genes involved in cellular adhesion and extracellular matrix remodeling correlates with poor survival of patients with renal cancer. *J Urol*. 2016;195(6):1892–902. <https://doi.org/10.1016/j.juro.2015.11.050>.
32. Chen S, Yang Y, He S, Lian M, Wang R, Fang J. Review of biomarkers for response to immunotherapy in HNSCC microenvironment. *Front Oncol*. 2023;13:1037884. <https://doi.org/10.3389/fonc.2023.1037884>.
33. Lin SC, Liao YC, Chen PM, et al. Periostin promotes ovarian cancer metastasis by enhancing M2 macrophages and cancer-associated fibroblasts via integrin-mediated NF- κ B and TGF- β 2 signaling. *J Biomed Sci*. 2022;29(1):109. <https://doi.org/10.1186/s12929-022-00888-x>.
34. Mao X, Xu J, Wang W, et al. Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives. *Mol Cancer*. 2021;20(1):131. <https://doi.org/10.1186/s12943-021-01428-1>.
35. Glabman RA, Choyke PL, Sato N. Cancer-associated fibroblasts: tumorigenicity and targeting for cancer therapy. *Cancers (Basel)*. 2022;14(16):3906. <https://doi.org/10.3390/cancers14163906>.
36. Ling W, Johnson SK, Mehdi SJ, et al. EDNRA-expressing mesenchymal cells are expanded in myeloma interstitial bone marrow and associated with disease progression. *Cancers (Basel)*. 2023;15(18):4519. <https://doi.org/10.3390/cancers15184519>.
37. Emmons TR, Giridharan T, Singel KL, et al. Mechanisms driving neutrophil-induced T-cell immunoparalysis in ovarian cancer. *Cancer Immunol Res*. 2021;9(7):790–810. <https://doi.org/10.1158/2326-6066.CIR-20-0922>.
38. Jiang C, et al. Role of HLA-DRB4 as a biomarker for endocrine toxicity and survival outcomes after immunotherapy in metastatic non-small cell lung cancer. *J Clin Oncol*. 2023;41:e21005–e21005.
39. Chen Z, Huang H, Wang Y, Zhan F, Quan Z. Identification of immune-related genes MSR1 and TLR7 in relation to macrophage and type-2 T-helper cells in osteosarcoma tumor micro-environments as anti-metastasis signatures. *Front Mol Biosci*. 2020;7:576298. <https://doi.org/10.3389/fmolb.2020.576298>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.