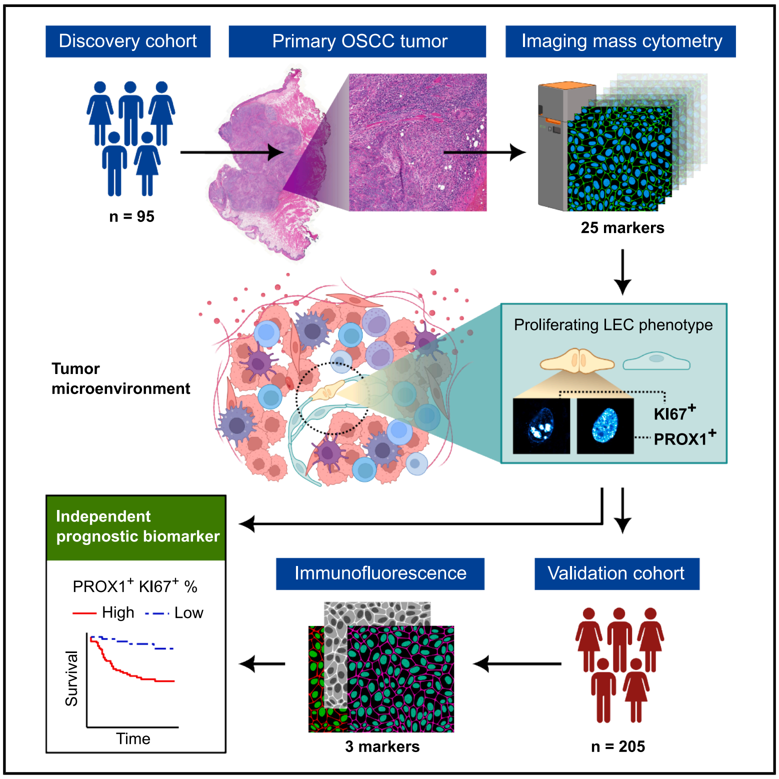
Cell Reports
Medicine

Article

Spatial single-cell analysis reveals tumor microenvironment signatures predictive of oral cavity cancer outcome

Graphical abstract



Authors

Joni Näsiaho, Linda Nissi, Sami Ventelä, Heikki Irjala, Marko Salmi

Correspondence

marko.salmi@utu.fi

In brief

Identifying those patients who progress to fatal disease after surgical treatment of early-stage OSCC is an unmet clinical need. Using TME proteomic single-cell analyses, Näsiaho et al. discover that proliferating LECs at the tumor's invasive margin serve as a clinically measurable independent adverse prognostic biomarker in two independent patient cohorts.

Highlights

- Spatial single-cell analyses of the tumor microenvironment in oral cavity cancer
- Proliferating lymphatic endothelial cells are an independent prognostic biomarker
- The biomarker is informative regardless of smoking status and in many treatment arms
- The $KI67^+PROX1^+$ biomarker outperforms TNM staging in early-stage oral cavity cancer

Näsiaho et al., 2026, Cell Reports Medicine 7, 102615
February 17, 2026
https://doi.org/10.1016/j.xcrm.2026.102615

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

Spatial single-cell analysis reveals tumor microenvironment signatures predictive of oral cavity cancer outcome

Joni Näsiaho,[1,2] **Linda Nissi,**[3] **Sami Ventelä,**[3,4] **Heikki Irjala,**[3] **and Marko Salmi**[1,2,5,6,*]
[1]MediCity Research Laboratory, University of Turku, Turku, Finland
[2]InFLAMES Research Flagship Center, University of Turku, Turku, Finland
[3]Department of Otorhinolaryngology, University of Turku and Turku University Hospital, Turku, Finland
[4]Turku Bioscience Centre, University of Turku and Åbo Akademi University, Turku, Finland
[5]Institute of Biomedicine, University of Turku, Turku, Finland
[6]Lead contact
*Correspondence: marko.salmi@utu.fi
https://doi.org/10.1016/j.xcrm.2026.102615

SUMMARY

Up to 20% of patients with early-stage oral squamous cell carcinoma (OSCC) develop postoperative relapse, but no parameters currently enable the identification of individuals with poor prognosis. Here, we report a spatially resolved single-cell analysis of the tumor microenvironment in OSCC. By analyzing >700,000 cells for 25 proteins using imaging mass cytometry, we discover leukocyte and endothelial cell phenotypes that are independent prognostic factors for survival. Most notably, the presence of proliferating lymphatic endothelial cells (KI67$^+$PROX1$^+$), which show distinct molecular signature in transcriptomic analyses, in the invasive tumor margin strongly predicts poor recurrence-free survival in early-stage OSCC. We validate our findings in an independent OSCC cohort using an easy-to-measure KI67$^+$PROX1$^+$ immunostaining biomarker assay and multivariate analyses of recurrence-free, disease-specific, and overall survival. Thus, our data highlight the role of tumor microenvironment in OSCC progression and its potential impact on treatment strategies.

INTRODUCTION

Oral squamous cell carcinoma (OSCC) is the most common type of malignancy affecting the oral cavity.[1,2] Despite the overall decrease in the incidence of tobacco-related tumors in recent years, OSCC incidence has steadily increased. Globally ~390,000 new cases and >188,000 deaths were reported in 2022. Among the OSCC locations, squamous cell carcinoma of the oral tongue (OTSCC) is the most common site of cancer in the oral cavity.[2,3] OTSCC incidence is rising in young adults. OTSCC is an aggressive cancer with a 5-year survival rate of ~60%,[4] and even with early-stage disease (i.e., T1-T2N0M0), 20% of patients will eventually relapse and face cancer-related death.[5]

OSCC is primarily treated with surgery including removal of the draining lymph nodes in all but T1 palate tumors.[2,6] Large tumor size, deep invasion, narrow resection margins, and metastatic seeding to cervical lymph nodes are the only clinically used unfavorable prognostic factors.[7–9] Despite intensive research, no molecular biomarker has been found to identify high-risk patients with early-stage OSCC.[10]

In addition to malignant cells, the tumor microenvironment (TME) is a critical regulator of tumor progression in many cancers, including OSCC.[11,12] Among tumor-infiltrating leukocytes, the subclass, activation status, and location of the cells are all decisive in determining their anti- versus pro-tumorigenic effects.[13,14] In general, the presence of cytotoxic T lymphocytes intermingling with tumor cells is associated with good outcomes, whereas the accumulation of immunosuppressive lymphocyte types, such as regulatory T cells (Tregs), or the exclusion of T cells from the tumor mass are poor prognostic signs.[15] The heterogeneous group of cancer-associated fibroblasts also exhibit both tumor-promoting and tumor-inhibiting properties linked to the production and remodeling of extracellular matrix, growth factor production, and leukocyte interaction.[16,17] Finally, blood and lymphatic vessels covered by endothelial cells (ECs) undergo neoangiogenesis in tumors and serve as portals for cancer cell metastasis and immune cell trafficking.[18–20]

We hypothesized that spatially resolved single-cell analyses of TME in OSCC might give new insights into tumor heterogeneity and reveal new prognostic biomarkers. We used imaging mass cytometry (IMC[21]) with a 25-antibody panel to characterize T lymphocytes, blood endothelial cells (BECs), lymphatic endothelial cells (LECs), and non-endothelial stromal cells (NESCs, mainly fibroblasts) in OSCC. We discovered that a proliferating LEC phenotype in the invasive tumor margin is a strong independent adverse prognostic factor of early-stage OSCC in two independent cohorts.

Table 1. Patient cohort characteristics

Characteristic[a]	Discovery cohort (n = 95)	Validation cohort (n = 205)	
	All patients	All patients	T1–T2
Patients included in the final analysis (n)	91	191	133
Age	67 (58, 74)	67 (58, 77)	67 (59, 75)
Gender			
Female	44 (48%)	102 (53%)	74 (56%)
Male	47 (52%)	89 (47%)	59 (44%)
Smoking status			
Non-smoker	39 (43%)	103 (54%)	71 (53%)
Active smoker	25 (27%)	63 (33%)	43 (32%)
Ex-smoker	25 (27%)	21 (11%)	15 (11%)
University Hospital District			
Helsinki	27 (30%)	0 (0%)	0 (0%)
Kuopio	20 (22%)	0 (0%)	0 (0%)
Oulu	14 (15%)	0 (0%)	0 (0%)
Tampere	15 (16%)	0 (0%)	0 (0%)
Turku	15 (16%)	191 (100%)	133 (100%)
pT			
1	40 (44%)	70 (37%)	70 (53%)
2	46 (51%)	63 (33%)	63 (47%)
3	5 (5.5%)	23 (12%)	0 (0%)
4	0 (0%)	35 (18%)	0 (0%)
pN			
0	72 (79%)	135 (71%)	112 (84%)
1	8 (8.8%)	20 (10%)	9 (6.8%)
2	4 (4.4%)	36 (19%)	12 (9.0%)
3	6 (6.6%)	0 (0%)	0 (0%)
OSCC subsite (ICD-10)			
Tongue (C02)	70 (77%)	83 (43%)	67 (50%)
Floor of mouth (C04)	11 (12%)	19 (9.9%)	10 (7.5%)
Gum (C03)	5 (5.5%)	30 (16%)	11 (8.3%)
Lip (C00)	0 (0%)	27 (14%)	23 (17%)
Other (C06)	5 (5.5%)	32 (17%)	22 (17%)
Surgery marginal (mm)	3.50 (2.00, 5.25)	NA	NA
Re-operation	15 (16%)	NA	NA
Treatment			
Surgery	66 (73%)	100 (52%)	90 (68%)
Surgery + RT or CRT	25 (27%)	80 (42%)	42 (32%)
Follow-up from treatment (months)	29 (17, 37)	55 (20, 60)	60 (37, 60)
Number of events			
Recurrence	13 (14%)	79 (41%)	44 (33%)
Death	17 (19%)	78 (41%)	38 (29%)

OSCC, oral squamous cell carcinoma; RT, radiotherapy; CRT, chemoradiotherapy.

[a] Median (interquartile range); n (%); NA, not assessed.

RESULTS

Geospatial analyses of the TME in early-stage OSCC

To analyze the TME in OSCC, we quantitatively explored the stromal landscape at a single-cell level using multiplexed IMC in a new prospective real-life patient cohort. The cohort consisted of all new early-stage (T1-T2N0M0) OSCC identified in Finland's five university hospitals in 2018–2019 (Table 1). Of the 95 patients, ∼50% were female, ∼45% were non-smokers, and ∼75% had OTSCC. None of the tumors showed regional or distant

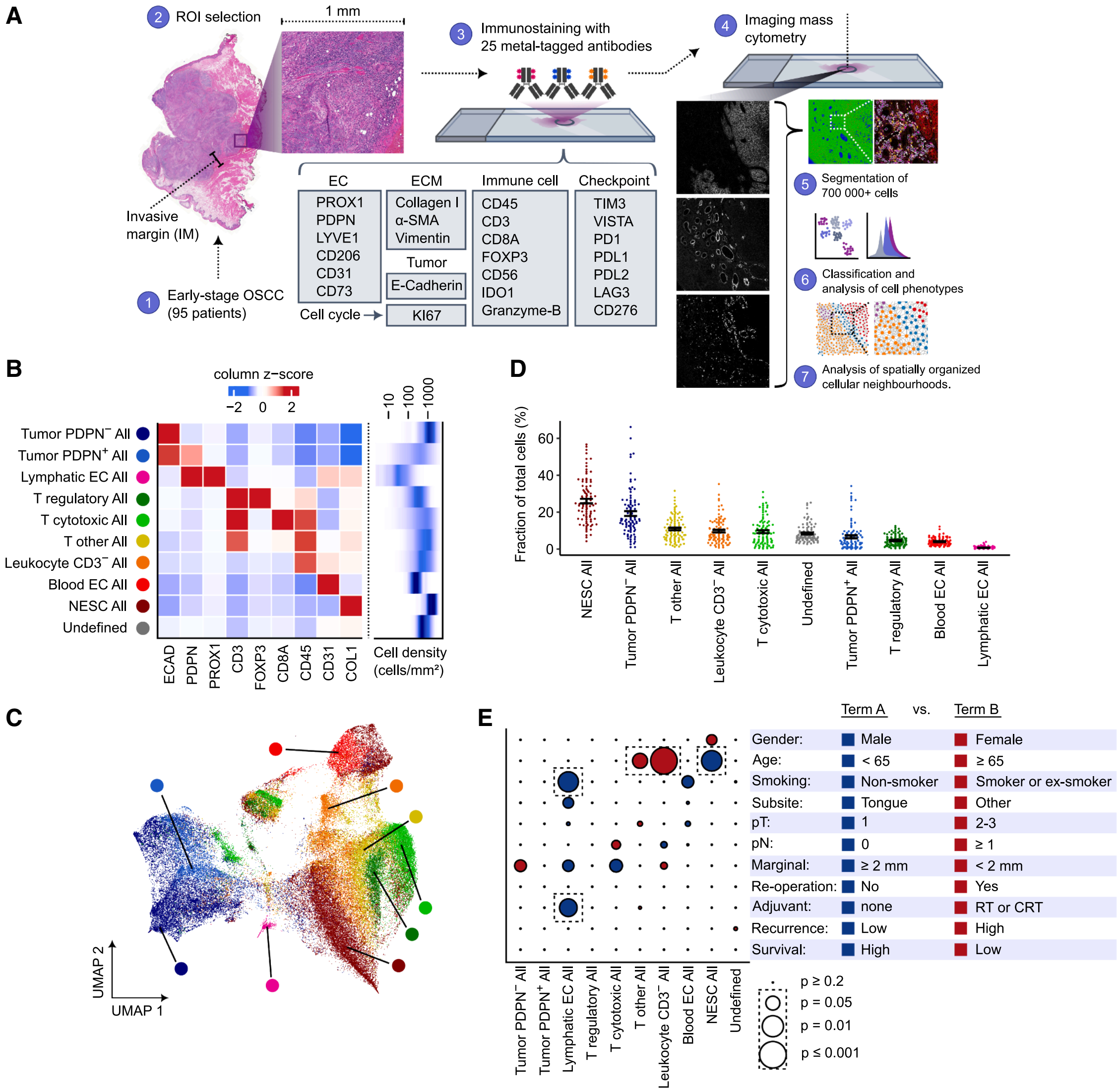


Figure 1. Characterization of the tumor microenvironment in early-stage OSCC

(A) Study setup in the discovery cohort outlining the selection of the tumor invasive margin (IM) area from OSCC samples, composition of the 25-antibody panel, and pipeline for multiplexed IMC staining and single-cell analyses.

(B) Heatmap of the main OSCC cell types, relative average expression of the indicated markers (*Z* score), and prevalence of the cell types (density/mm^2).

(C) UMAP plot of the main cell types (a random subset of 80,000 cells; same color code as in B).

(D) Frequency of the main cell types in OSCC. Each dot represents a single sample, and the lines indicate mean ± standard error.

(E) Cell frequency comparisons between the cell types and indicated clinical variables. The bubble color indicates the clinical subgroup (term A or B) with a higher cell type frequency, and the bubble size indicates the *p* value. Statistically significant differences are highlighted by gray dotted lines (Student's *t* test, except log rank test for analyzing the recurrence and survival). RT, radiotherapy; CRT, chemoradiotherapy.

See also Figures S1 and S2; Tables S1 and S2.

metastasis on radiological examination, and they were operated with curative intent. In subsequent pathologic examination, 40 T1, 46 T2, and 5 T3 stage cases were identified. Of all patients, ∼80% were free of lymph node involvement (pN0) (Table 1), and lymph node status was the only prognostic characteristic for recurrence-free survival (RFS) in this cohort according to univariate and multivariate Cox proportional hazard models (Table S1).

We used whole formalin-fixed paraffin-embedded (FFPE) sections from pathologist-verified primary OSCC tumors for IMC. We stained sections with a panel of 25 metal-labeled antibodies, which were selected to comprehensively phenotype ECs, fibroblasts, and T lymphocytes with an emphasis on immunosuppressive and checkpoint markers (Figure 1A; Table S2). We verified the performance of 17 commercially conjugated and 5 in-house conjugated antibodies based on their expected staining patterns in tumor and lymph node sections (Figures S1A–S1C). After IMC imaging of one 1 mm × 1 mm pre-determined region of interest (ROI) at the invasive margin (IM) in each patient, we used nuclear, cytoplasmic, and cell surface markers to generate Ilastik cell probability images and subsequent cellular segmentation masks[22] to quantify multiplexed markers at a single-cell resolution (Figure 1A).

Spatial cellular neighborhoods in OSCC

We identified E-cadherin$^+$ (ECAD$^+$) tumor cells, CD3$^+$ T lymphocytes, CD45$^+$CD3$^-$ non-T cell leukocytes, and PROX1$^+$ LECs in OSCC samples using supervised cell labeling and a random forest classifier (Figures 1B, S2A, and S2B). The ECAD$^+$ tumor cells were subdivided into podoplanin (PDPN)-positive and negative cells, and the CD45$^+$CD3$^+$ immune cells were subcategorized into CD8$^+$ T cytotoxic, FoxP3$^+$ T regulatory, and CD3$^+$CD8$^-$FoxP3$^-$ other T cells. Using further clustering with all markers, two additional major cell populations, CD31^{high}PROX1$^-$ BECs and COL1$^+$ NESCs, were identified, while the remaining cells were left undefined.

When using uniform manifold approximation and projection (UMAP) for dimensional reduction, tumor cells, LECs, and BECs formed clearly separate cell clusters, whereas NESCs and leukocyte clusters were more intermingled with other cell types (Figure 1C). We identified 25.9% ± 12.8% of the sample cells as tumor cells, 34.8% ± 17.0% as immune cells, and 30.8% ± 12.6% as stromal cells (endothelial or non-endothelial), and 8.5% ± 5.0% of cells could not be assigned with our panel. NESCs and PDPN$^-$ tumor cells were the most frequent cells in the samples, whereas BECs (n = 29,775) and LECs (n = 5,302) represented the rarest cell types (Figures 1B and 1D). No cell type correlated with tumor size, lymph node metastasis, disease recurrence, or survival (Figure 1E).

The spatial distribution of cells at the tumor IM affects the propensity of tumor cells to invade locally, form metastasis via lymphatic and blood vessels, and interact with immune cells.[11,23] We therefore analyzed interaction (attraction) and avoidance behaviors between the different cell types[24,25] in OTSCC and other types of OSCC (Figures 2A and S3A). In most samples of both patient groups, PDPN$^-$ and PDPN$^+$ tumor cells showed significant interactions and formed a distinct tumor compartment, whereas they avoided the main immune and stromal cell types. We found that PDPN$^+$ tumor cells avoided LECs less strongly than PDPN$^-$ tumor cells, which is in line with PDPN expression mainly at the edge of individual tumor cell patches (Figure S3B). All four main leukocyte subtypes preferentially interacted with each other but avoided tumor cells (Figure 2A). Among the main leukocyte types, LECs and BECs interacted most with CD8$^-$FoxP3$^-$ other T cells, which were also furthest away from the tumor cells (Figures 2A and 2B). We found only a few differences in the interaction patterns when comparing the behavior of OTSCC and other OSCC cancers (Figures 2A and S3C).

We identified two tumor-enriched spatial neighborhoods at the IM using an unsupervised local indicator of spatial association method[26] (Figures 2C, 2D, and S3D). These spatial regions were designated as "IM tumor/central" and "IM tumor/border," according to their localization closer to or farther away from the tumor-stroma interface, respectively. Based on the most enriched cell types, the three non-tumor cell-enriched spatial areas were classified as "IM stroma-rich region," "IM pan-leukocyte hotspot," and "IM T cell hotspot." The proportion of the spatial areas showed marked interindividual variation, e.g., the IM stroma-rich region occupied 10.2%–52.5% of the sample area (24.9% ± 8.5%, mean ± SD; Figure 2D). However, none of the five spatial regions correlated with tumor size, lymph node metastasis, recurrence, or survival (Figure 2E).

Intratumoral leukocyte phenotypes predict survival in OSCC

We next analyzed the value of refined phenotypes of tumor cells, NESCs, and leukocytes in predicting survival following standard primary treatment of early-stage OSCC. We identified six tumor cell and five NESC subpopulations, but none of them showed any association with survival in Kaplan-Meier analyses (Figures S4A–S4E).

We subclassified the three main T lymphocyte populations (cytotoxic, regulatory, and other T cells) into nine subpopulations using unsupervised clustering together with manually defined localization within or outside the tumor at the IM (Figure 3A). The three main T cell populations were not associated with survival in OSCC (Figure S4C), but patients with higher-than-median levels of GZYB$^+$ cytotoxic CD8$^+$ T cells showed shorter RFS and overall survival (OS) in Kaplan-Meier plots (Figure 3B). Moreover, higher intratumoral T cell counts of the CD8$^+$ or FoxP3$^+$ phenotype were associated (p = 0.003 and p = 0.005, respectively) with improved OS (Figure 3C). The other T cell or non-T cell subpopulations were not associated with survival (Figures S4C, S4F, and S4G).

Using Cox proportional hazard models adjusted for the prognostic factors age, tumor size (pT), and nodal involvement (pN), we discovered that a high CD8$^+$GZYB$^+$ phenotype independently predicts impaired RFS (hazard ratio [HR]: 6.7, 95% confidence interval [CI]: 1.7–26.0, p = 0.006) and impaired OS (HR: 3.6 [95% CI: 1.2–10.9], p = 0.02; Figure 3D). In contrast, a high frequency of intratumoral CD8$^+$ T lymphocytes was independently associated with longer RFS (HR: 0.25 [95% CI: 0.07–0.94], p = 0.04) and longer OS (HR: 0.10 [95% CI: 0.02–0.40], p = 0.001). A high frequency of intratumoral Tregs also predicted longer OS in this multivariate Cox model (Figure 3D; Table S3). The frequency of FoxP3$^+$ and CD8$^+$ cells was similar in "IM tumor" spatial areas, whereas CD8$^+$ cells were predominant in the three "IM stroma" areas (Figure 3E). The FoxP3:CD8 cell ratio had no prognostic value for survival (Figure 3F). T cells expressing exhaustion markers PD1, TIM3, and VISTA[27] did not predict survival in whole IM samples, IM T cell hotspots (where the exhausted cell types were enriched), or in IM tumor regions (Figures 3G and S4H–S4M).

IM T cell hotspots were enriched in smokers compared to non-smokers (Figure 2E), which led us to further study the associations between smoking and the immune landscape. Refined cellular neighborhood analyses showed that few cell-cell interactions differed between non-smokers and smokers (Figures S5A and S5B). However, intratumoral T cells (Tc IT and Treg IT) were less likely to avoid LECs in smokers than in non-smokers. We also found that smokers had higher frequencies of all measured leukocyte types in the IM T cell hotspot areas but lower frequencies of CD3$^-$ leukocytes in the IM tumor/border niche and "IM stroma-rich" areas (Figure S5C). The prognostic value of different immune cells in both subgroups remained quite similar to the pooled analyses, but the only statistically significant association was the

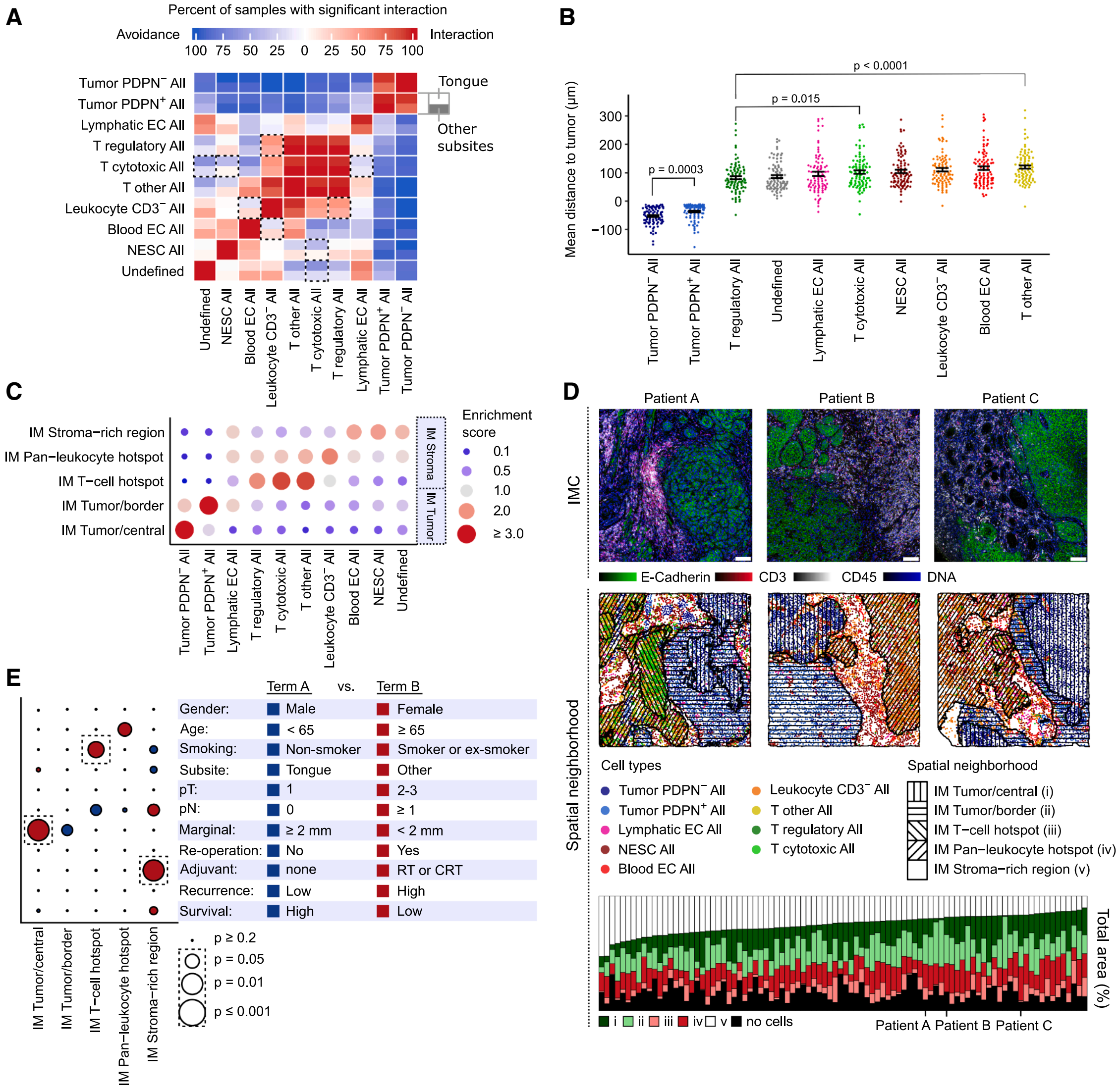


Figure 2. Spatial cellular neighborhoods in OSCC

(A) Heatmap of pairwise cellular interaction-avoidance scores (percent of samples with significant interaction [red] or avoidance [blue]). In each grid, the upper and lower rows represent OTSCC (tongue) and other OSCC types (other subsites), respectively. Statistically significant differences between OTSCC and other OSCC tumor types are highlighted by gray dotted lines.

(B) The distance between different cell types and the nearest tumor cell. Each dot represents a single sample, and the lines indicate mean ± standard error.

(C) Distribution of different cell types within the five spatial neighborhood regions in the IM of the OSCC samples. The enrichment scores were calculated relative to the expected random distribution.

(D) Distribution of spatial regions in the OSCC TME. Uppermost row: representative IMC raw images of E-cadherin, CD3, CD45, and DAPI (DNA) staining in three patients A–C with OSCC. Middle row: localization of individual cell types (the color code below the image) within the five spatial regions (i–v, fill-in pattern shown below the image) in patients A–C. Bottom row: distribution of the spatial regions i–v (as defined above the image) as a percentage of the total area in all individual patient samples (*n* = 91).

(E) Frequency comparisons of the OSCC regions with the indicated clinical variables. The bubble color indicates the clinical subgroup with a higher spatial area frequency, and the bubble size indicates the *p* value. Statistically significant differences are highlighted by gray dotted lines (Student's *t* test, except log rank test for analyzing the recurrence and survival). RT, radiotherapy; CRT, chemoradiotherapy.

See also Figure S3.

better OS in non-smokers with high numbers of intratumoral cytotoxic T cells (Figures S5D–S5F).

Hence, in early-stage OSCC, the presence of intratumoral cytotoxic T cells or Tregs at the IM predicts improved survival, whereas phenotypically exhausted T cells do not correlate with prognosis.

Proliferating LECs independently predict poor outcome in OSCC

We discovered that tumor IM LECs, defined as PROX1$^+$ cells,[28] formed five clusters based on the expression of PDPN, LYVE1, CD31, and KI67 (Figure 4A). Total LEC numbers were not

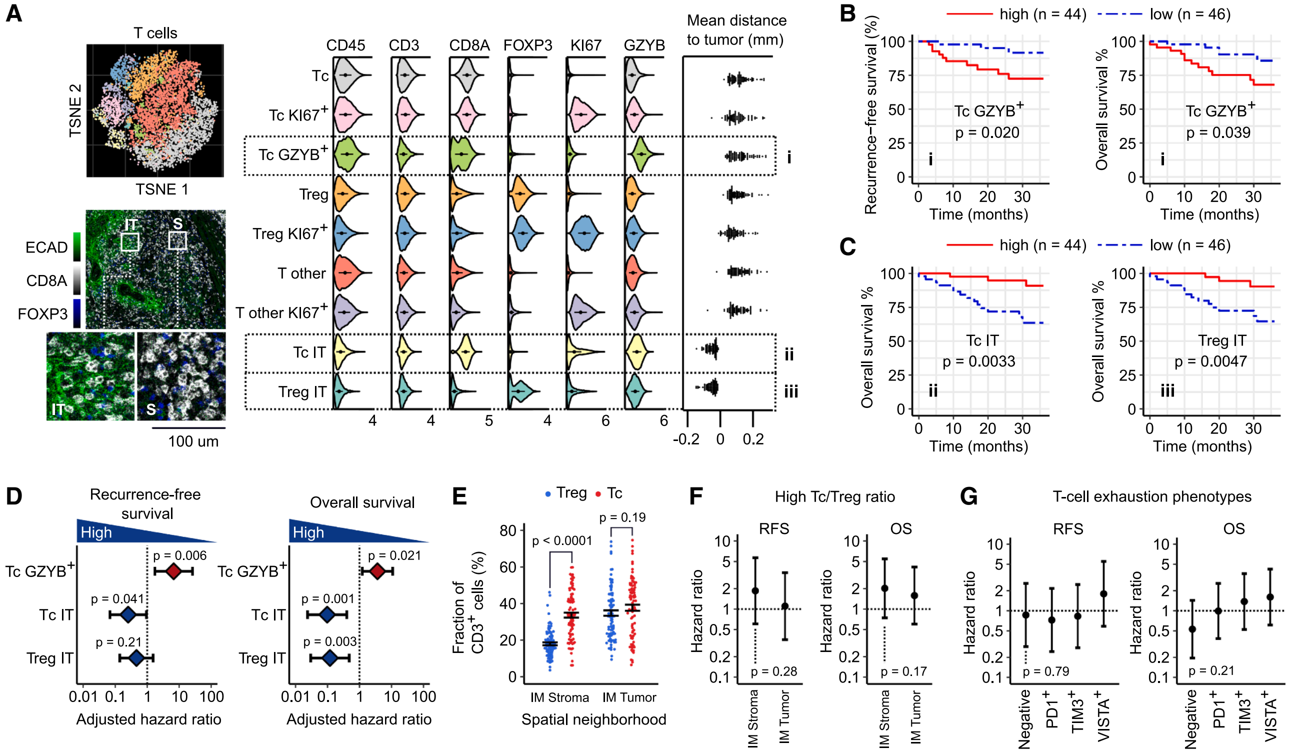


Figure 3. Phenotypic heterogeneity and prognostic value of T lymphocytes in OSCC

(A) T-distributed stochastic neighbor embedding (t-SNE) visualization of T cells (randomly sampled 5,000 cells; colored by the identified T cell clusters), marker expression, and mean distance to the tumor. A representative IMC image showing the intratumoral (IT) and stromal (S) localization of $CD8^+$ and $FOXP3^+$ cells at the IM.

(B and C) Kaplan-Meier analyses showing RFS and OS based on the frequency of a given T cell subpopulation (Tc $GZYB^+$ cells, B, and intratumoral $CD8^+$ [TcIT] and $FOXP3^+$ [Treg IT] cells, C) above (high) or below (low) the median frequency of that subpopulation.

(D) Cox proportional hazard models (HR and 95% CI) adjusted for age, pT, and pN with the indicated T cell subpopulations for RFS and OS.

(E) The proportions of cytotoxic $CD8^+$ (Tc) and $FOXP3^+$ (Treg) T cells (out of all T cells) within the indicated spatial areas.

(F) Univariate Cox proportional hazard model of high cytotoxic T cell-to-Treg ratio in the indicated spatial areas for RFS and OS.

(G) Univariate Cox proportional hazard model of exhausted ($PD1^+$, $TIM3^+$, and $VISTA^+$) T cells and non-exhausted T cells (negative) for RFS and OS.

See also Figures S4 and S5; Table S3.

associated with survival in OSCC (Figures 1E and S4C; Table S3). However, high numbers of $LYVE1^-$ LECs were associated with favorable OS (Figure 4B), but this LEC subpopulation did not have an independent predictive value (Figure 4C). In contrast, a higher-than-median percentage of proliferating $KI67^+$ LECs was strongly associated with poor RFS ($p = 0.002$) and OS ($p = 0.002$) (Figure 4D). Notably, at 36 months, we observed OSCC recurrence in only 6% of patients with a low $KI67^+$ phenotype compared with 32% of patients with a high $KI67^+$ phenotype. Single-cell analyses further revealed that only $KI67^+$ LECs, but not $KI67^-$ LECs, had value in predicting recurrence (Figure 4E; Table S3).

Using a fully adjusted (age, pT, and pN) Cox model, we discovered that $KI67^+$ LECs were a strong independent factor for predicting RFS (HR: 8.3 [95% CI: 1.8–38.9], $p = 0.008$) and OS (HR: 5.9 [95% CI: 1.6–21.1], $p = 0.007$) (Figure 4F; Table S3). In fact, only $KI67^+$ LECs and lymph node metastasis status were predictive of RFS, whereas $KI67^+$ LECs were the only marker predicting poor OS in these patients (Figure 4F). The prognostic power of proliferating LECs remained intact, even when smoking was included in the adjusted Cox model (Figure S6A).

PROX1 and KI67 co-expression in the LEC nucleus allowed us to determine cell phenotypes with high confidence. We found stronger PROX1 signals in non-nucleolar areas than within nucleoli, whereas KI67 expression was high in nucleoli and weaker at the periphery of the nucleus (Figure 4G). The proportion of proliferating LECs ($KI67^+$ LECs/all LECs) varied substantially (from 0% to 50%; 5.4% ± 9.5%, mean ± SD) among patients (Figures 4H and S6B). When we divided lymphatic vessel profiles into those with none, one, or two or more $KI67^+PROX1^+$ cells (assigned as LVp0, LVp1, and LVp2, respectively), we observed that lymphatic vessels with ≥ 2 $KI67^+$ LECs were highly enriched at the spatial region IM tumor/border (Figure 4I). The lymphatic vessels in IM tumor areas with no proliferating cells (LVp0), representing on average 25% ± 13% (mean ± SD) of all lymphatics, had absolutely no correlation with RFS (Figure 4I). In stark contrast, a high number of lymphatic vessels with one or more $KI67^+$ LECs in IM tumor areas associated very strongly ($p = 0.0003$) with poor prognosis (Figure 4I).

In sensitivity analyses, $KI67^+$ lymphatic vessels (LVp1-2) were associated with disease progression in both patients with no

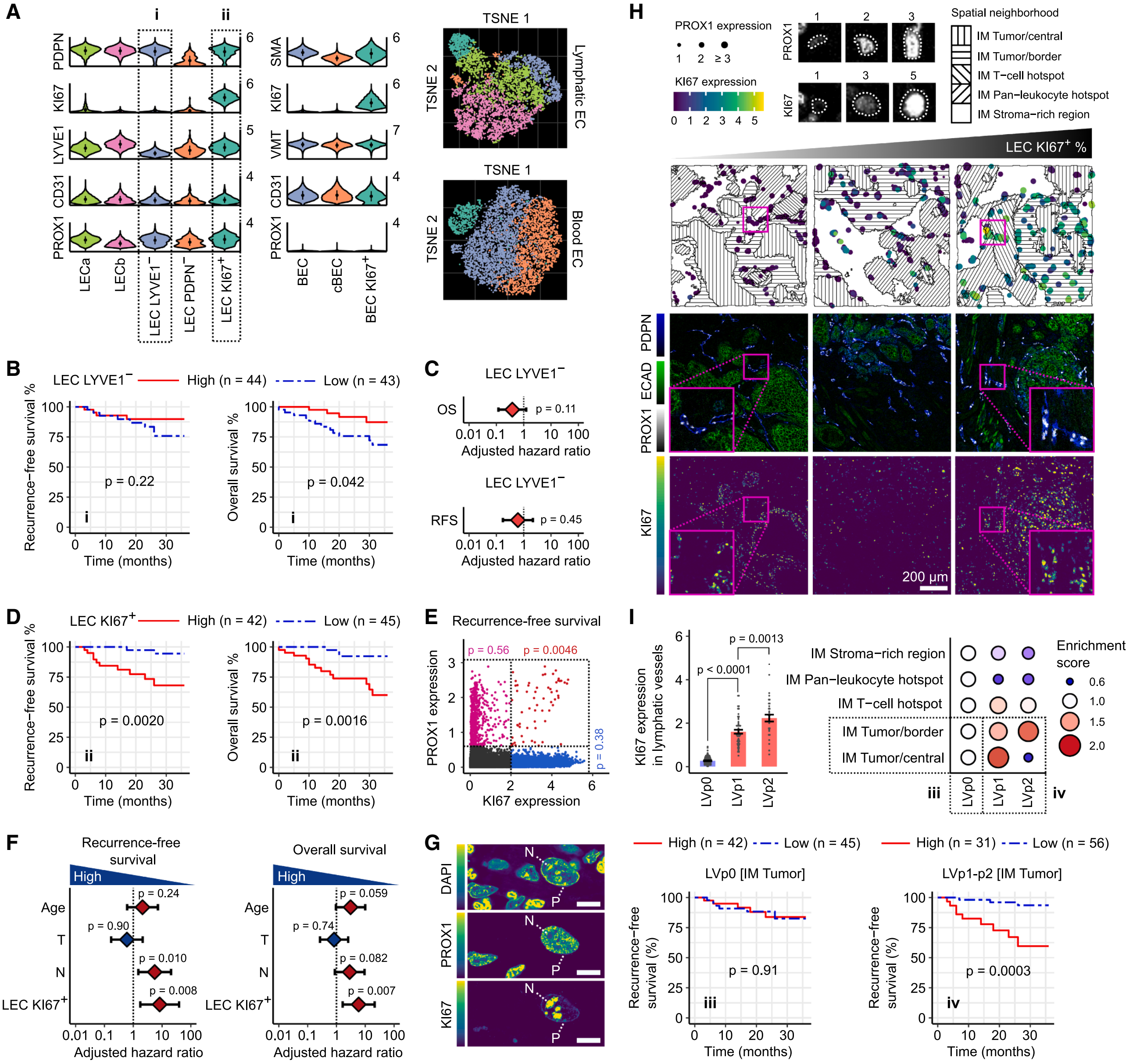


Figure 4. Phenotypic heterogeneity and prognostic value of endothelial cells in OSCC

(A) Violin plots showing the expression of the indicated markers in endothelial subpopulations. t-SNE visualization of endothelial cells (randomly sampled 5,000 cells; colored by the identified clusters) shows the distribution of the LEC and BEC subpopulations.

(B) Kaplan-Meier analyses of survival in patients with a frequency of LEC LYVE1$^-$ cells above (high) or below (low) the median frequency.

(C) Cox proportional hazard models (adjusted for age, pT, and pN) showing RFS and OS for the LEC LYVE1$^-$ high patients.

(D) Kaplan-Meier plots for KI67$^+$ high and low LEC phenotypes for RFS and OS.

(E) Single-cell scatterplots showing the PROX1$^+$KI67$^-$, PROX1$^+$KI67$^+$, and PROX1$^-$Ki67$^+$ populations and their prognostic value in RFS (log rank test).

(F) Multivariate Cox proportional hazard model (HR and 95% CI) adjusted for age, pT, pN, and KI67$^+$ LECs in predicting RFS and OS. Note the logarithmic scale.

(G) High-resolution confocal microscope image showing PROX1 and KI67 expression in a representative OSCC sample. N, nucleolus; P, perinucleolar area.

(H) Distribution and intensity of PROX1 and KI67 expression in the five spatial areas of OSCC tumors. Three representative samples with none, moderate, and high frequencies of dividing KI67$^+$PROX1$^+$ LECs.

(I) Prognostic value of lymphatic vessels containing proliferating LECs. Distribution of lymphatic vessels, defined as vessel profiles containing zero (LVp0), one (LVp1), or two or more (LVp2) KI67$^+$PROX1$^+$ LECs within the spatial neighborhood regions of the OSCC samples. The enrichment scores are presented relative to the LVp0 population. Kaplan-Meier analyses showing the association between LVp0 and LVp1-2 vessel frequency at IM tumor spatial area and patient survival.

See also Figures S4 and S6; Tables S3 and S4.

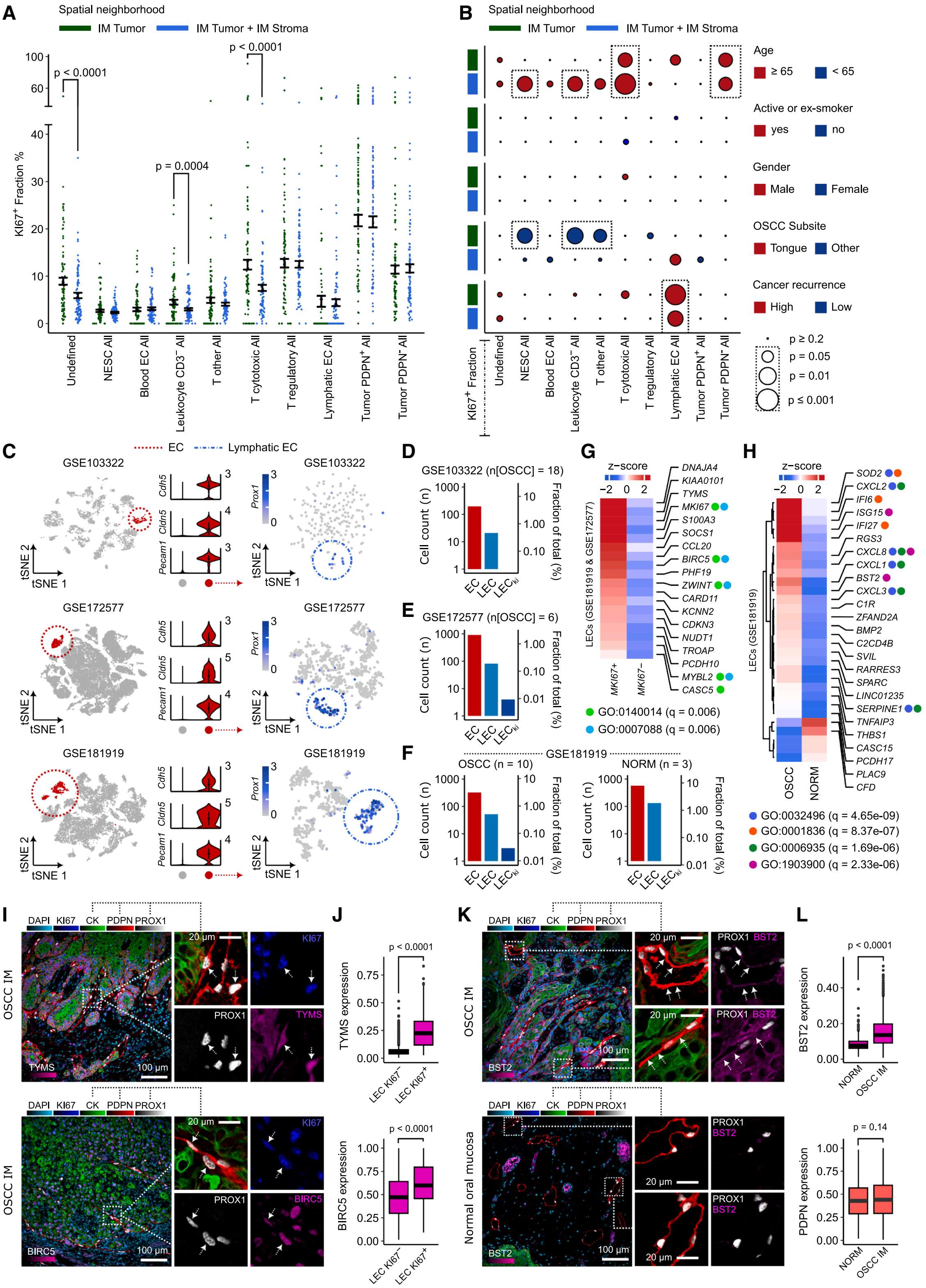


Figure 5. Proliferating KI67$^+$ cell populations in OSCC

(A) Percentage of KI67$^+$ cells within each cell type in IM tumor (green) spatial areas and in the whole sample ("All", blue). Paired comparison with Wilcoxon-signed rank test.

(legend continued on next page)

nodal involvement (pN = 0, $p = 0.007$) and in the few patients with locally spread tumors (pN+) (Figure S6C). Nonproliferating LECs had no value in predicting survival in lymph node-negative or -positive cases (Figure S6C). The proliferating lymphatic vessels did not differ from nonproliferating ones in terms of PDPN, LYVE1, CD31, and CD73 expression, but they did express higher levels of VMT, CD206, and PROX1 (Figure S6D). Moreover, LECs expressed PDL1 (which can contribute to local immune suppression[29]) particularly in the KI67$^+$ and PDPN$^-$ subpopulations (Figure S2A). PDL1 expression in LECs correlated with the density of tumor-infiltrating leukocytes (Figure S6E) but was not associated with survival (Figure S6E).

Additional Kaplan-Meier analyses showed that patients positive for both independent risk factors for poor RFS (high KI67$^+$ LECs and pN+) had markedly adverse outcome compared with those with only one or no risk factor ($p < 0.0001$), and 70.0% of the high-risk group had experienced a recurrence in 3 years (Figure S6F). Cox analyses further revealed that the combination score predicted RFS (HR: 19.0 [95% CI: 5–102]) and OS (HR: 9.5 [95% CI: 2.3–40.2]) better than either of its components separately (Figure S6F; Table S4). Taken together, proliferating LECs, but not total LECs or other LEC subtypes, at the IM have strong independent prognostic value for poor outcome in early-stage OSCC.

Proliferating LECs have a distinct molecular signature in OSCC

Many non-LEC cells harbored a sizable KI67$^+$ subpopulation in OSCC (Figure 5A). We found the highest proliferative fraction (25.7% ± 14.0%) in PDPN$^+$ tumor cells, and PDPN$^-$ tumor cells, cytotoxic T cells, and Tregs had 8%–13% KI67$^+$ cells. Of note, BECs also contained a clear KI67$^+$ subpopulation (3.4% ± 2.7%) (Figure 5A). Among cytotoxic T cells and CD3$^-$ leukocytes (and undefined cells), the proliferating cells were enriched to IM tumor, whereas among other cell types (including LECs), the percentage of proliferating cells was similar in this spatial area and in the whole IM sample (Figure 5A). The fraction of proliferating LECs did not correlate with age, gender, smoking status, or tumor subsite, whereas in many other cell types, proliferation correlated positively to age (Figure 5B). Strikingly, only KI67$^+$ LECs, but not KI67$^+$ BECs or any other KI67$^+$ cell type, correlated to cancer recurrence (Figure 5B) or predicted OS or RFS (Figure S4C; Table S3).

OSCC has been studied using single-cell RNA sequencing (scRNA-seq) in search of biomarkers,[30–32] but LECs have not been analyzed in those studies. We identified a small but clear endothelial cell cluster (197–909 cells) in each dataset (altogether 34 patients and 65,141 cells) using established pan-endothelial markers *PECAM1*, *CLDN5*, and *CDH5* (Figure 5C). When reclustering the ECs at a higher resolution, we recognized a well-defined *PROX1*$^+$ LEC subpopulation, and in two out of three datasets, we found a few *MKI67*$^+$ *PROX1*$^+$ cells (Figures 5C–5F). Several genes, such as *DNAJ4* (regulating heat shock response and endothelial migration[33]), *TYMS* (involved in DNA replication[34]), *SOCS1* (regulating JAK/STAT signaling[35]), and *CCL20* (supporting chemotaxis of Tregs and other lymphocytes[36] and angiogenesis[37]), were enriched in proliferating LECs compared to nonproliferating LECs (Figure 5G). Of note, we observed no proliferating LECs in three normal samples from oral mucosa (Figure 5F). Molecules involved in neutrophil chemotaxis, such as chemokines *CXCL1*, *CXCL2*, *CXCL3*, and *CXCL8*, and interferon-inducible molecules (e.g., *IFI6*, *ISG15*, and *BST2*), were enriched in OSCC LECs compared to normal LECs (Figure 5H). On the other hand, normal LECs expressed more *TNFAIP3*, a terminator of NF-κB activation and inhibitor of angiogenesis,[38] and *THBS1*, a cell-surface adhesion protein and an inhibitor of lymphangiogenesis.[39]

We validated a set of identified marker candidates for proliferating and tumor-associated LECs using immunostaining and cell segmentation. In LECs, TYMS outlined the nucleus and whole cytoplasm of proliferating KI67$^+$ LECs, whereas it was absent from nonproliferating LECs (Figures 5I and 5J). BIRC5 protein was enriched in the nuclei of proliferating LECs (Figures 5I and 5J). When comparing normal and OSCC LECs, we found increased BST2 expression in tumor-associated LECs (Figures 5K and 5L), whereas no difference was observed in the expression of the common LEC marker PDPN. Thus, protein staining faithfully reproduced the scRNA-seq data with the two

(B) Cell frequency comparisons between the KI67$^+$ proliferating cells and the indicated clinical variables. The bubble color indicates the clinical subgroup with the higher frequency of the proliferating cells, and the bubble size shows the p value. The analyses are done separately for the IM tumor areas and for the whole sample. Statistically significant differences are highlighted by gray dotted lines (Mann-Whitney U test, except log rank test for analyzing the recurrence).

(C) t-SNE plots (left) visualizing three human OSCC scRNA-seq datasets acquired from the Gene Expression Omnibus database, with ECs highlighted by a red circle. Violin plots (middle) show the expression of the indicated EC markers. t-SNE plots (right) show PROX1 expression in ECs. LECs are highlighted with a blue circle.

(D–F) Bar plots showing total scRNA-seq cell counts and fractions of EC populations in OSCC (GSE103322, D, and GSE172577, E) and in OSCC and normal oral mucosa (NORM) (GSE181919, F). LECki = *MKI67*$^+$ LECs.

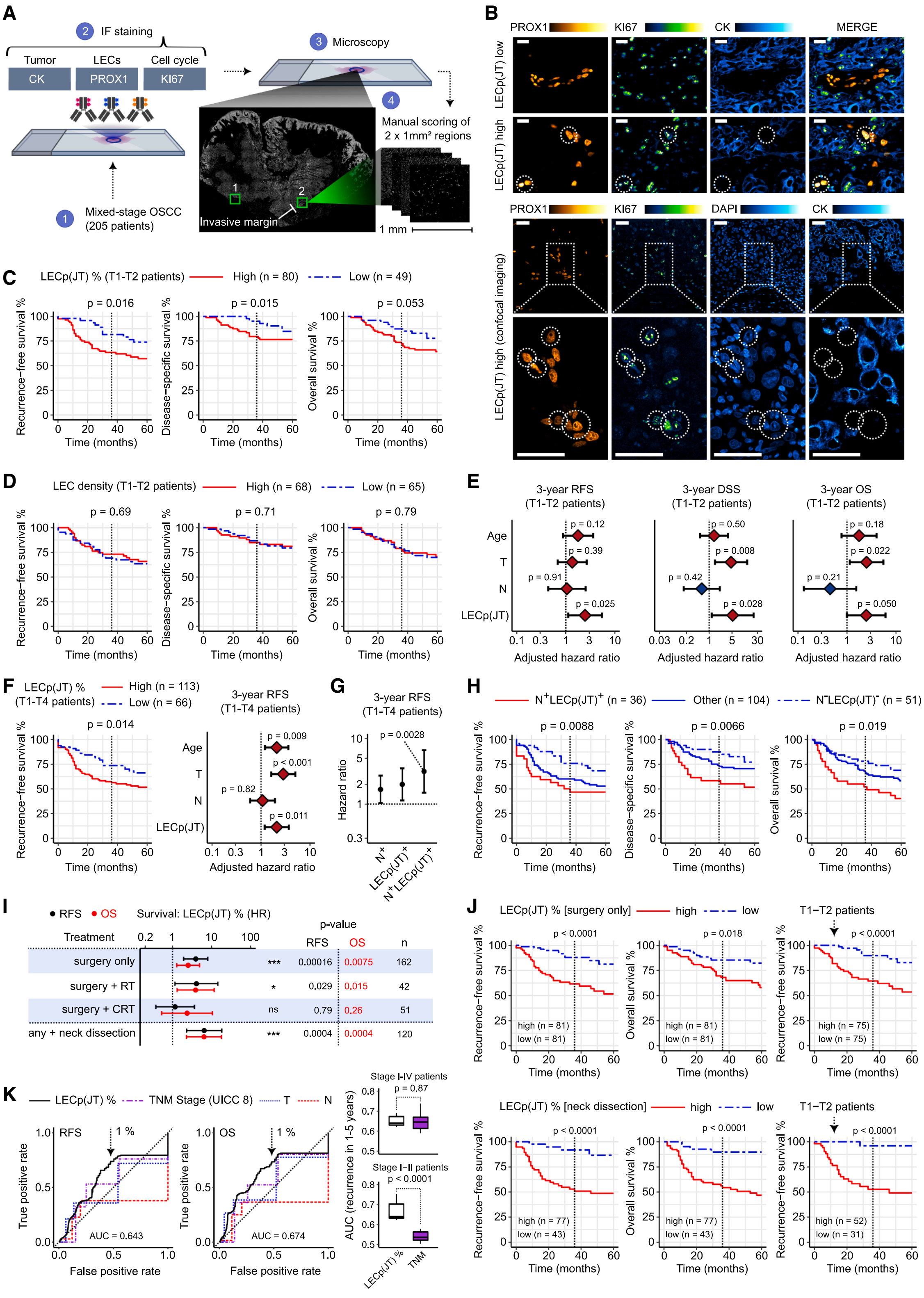
(G) Heatmap of the scaled average expression of 20 differentially expressed genes between *MKI67*$^+$ and *MKI67*$^-$ LECs in combined analysis for datasets GSE172577 and GSE181919. Significantly enriched (over-representation analysis [ORA]) Gene Ontology (GO) pathways are marked with colored circles for indicated marker. GO:0140014, mitotic nuclear division; GO:0007088, regulation of mitotic nuclear division. q value, false discovery rate-adjusted p value.

(H) Heatmap of the scaled average expression of 30 differentially expressed genes between LECs in OSCC and NORM in dataset GSE181919. Colored circles represent significantly enriched GO pathways for the indicated marker. GO:0032496, response to lipopolysaccharide; GO:0001836, release of cytochrome *c* from mitochondria; GO:0006935, chemotaxis; GO1903900, regulation of viral life cycle. q value, false discovery rate-adjusted p value.

(I and J) Expression (I) and quantification (J) of TYMS and BIRC5 in OSCC. The samples were co-stained for PROX1, PDPN, CK, and KI67. Arrows indicate proliferating LECs.

(K and L) Expression (K) and quantification (L) of BST2 and PDPN in OSCC and normal oral mucosa LECs. The samples were co-stained for PROX1, CK, and KI67. Arrows indicate BST2$^+$ LECs in the tumor.

Data in (I)–(L) are from 3 to 5 patients (6–8 regions/patient and 5,569 LECs). Mann-Whitney U test.



(legend on next page)

tested proliferation-associated molecules (TYMS and BIRC5) and one proliferation non-associated molecule (BST2).

Collectively, our data indicate that only LECs are associated with OSCC outcome among KI67$^+$ cells and that LECs display new molecular characteristics at the mRNA and protein level that are useful when differentiating proliferating and nonproliferating LECs in OSCC and between tumor-associated and normal LECs.

Proliferating LECs are a clinically useful biomarker of OSCC

To determine whether LEC proliferation status is a useful biomarker in clinical settings, we developed immunofluorescence staining for PROX1, KI67, and pan-cytokeratin using wide-field microscopy and defined optimal cutoff points for percentage of proliferating LECs and for the minimal LEC density to obtain reliable phenotype assignment using the discovery series (Figures S6G–S6I).

For the validation experiments, we stained and analyzed an independent series of 205 OSCC whole-section tumors (Figures 6A and 6B). Samples of 191 patients, of which 133 had T1–T2 tumors, contained adequate tumor-stroma interface for analyses (Table 1). The clinicopathological characteristics of patients with early-stage OSCC in the discovery and validation series were statistically similar, with the exception of geographical area distribution, OTSCC versus non-OTSCC ratio, and follow-up time (Table 1). For quantitative analyses, we chose two 1-mm^2 regions from the IM and determined the mean numbers of juxtatumoral (defined as ~50 μm from the nearest tumor cells) proliferating (KI67$^+$PROX1$^+$) and overall (PROX1$^+$) LECs using manual scoring (Figures 6A and 6B).

We discovered that a high percentage of proliferating juxtatumoral LECs was associated with poor RFS and disease-specific survival (DSS) in patients with T1–T2 OSCC tumors in the validation cohort (Figure 6C). The overall LEC density did not correlate with survival (Figure 6D). In multivariate Cox analyses, KI67$^+$PROX1$^+$ cells, but not age, tumor grade, or lymph node metastasis, independently predicted 3-year RFS (Figure 6E; Table S4). Proliferating LECs (and T grade) independently predicted poor 3-year DSS and OS among these patients. Moreover, KI67$^+$PROX1$^+$ LECs were an independent prognostic factor for 5-year RFS in early-stage OSCC (Figure S7A; Table S5).

When analyzing all patients (T1–T4) of the validation cohort, we found that a high number of proliferating juxtatumoral LECs were associated with a poor 3-year RFS (Figure 6F). In multivariate models, KI67$^+$ LECs remained an independent prognostic factor for 3-year RFS, DSS, and OS and for 5-year RFS and OS (Figures 6F and S7B; Table S5). The overall LEC density showed no prognostic value among all patients (Figure S7C). Finally, when counting the combined risk score using LEC proliferation and nodal status (as in the discovery cohort) for the whole validation cohort (T1–T4), we observed that it predicted decreased 3-year and 5-year RFS, DSS, and OS markedly better than either parameter alone (Figures 6G and 6H; Table S4).

In sensitivity analyses, we discovered that when the patients were categorized into high- and low-LEC proliferation groups based on the maximal number of KI67$^+$PROX1$^+$ cells (i.e., using proliferating LEC data only from the IM spot having the higher percentage), the high group both among T1–T2 and all patients showed worse RFS, DSS, and OS (Figures S7D and S7E). KI67$^+$PROX1$^+$ LECs independently predicted poor 3-year and 5-year RFS and OS in T1–T2 patients in the validation cohort even when smoking was included in the Cox models (Figure S7F).

We also used deep learning-based cell segmentation to define proliferating LECs at IM in the validation series and found an excellent correlation between the manual and automated methods in terms of both LEC density and LEC proliferation quantitation (Figure S7G). Moreover, patients belonging to the high KI67$^+$PROX1$^+$ group based on segmentation analyses had impaired RFS and OS (Figure S7H). We further performed cell segmentation analyses for 20 whole sections with high proliferating LEC counts. Consistent with the other data, we observed

Figure 6. Validation of the KI67$^+$PROX1$^+$ biomarker in an independent OSCC cohort and its performance in different treatment arms and against TNM staging

(A) Validation study setup outlining the selection of the three-antibody panel and the pipeline for conventional immunofluorescence analyses (used for B–H).

(B) Wide-field (upper) and confocal (lower) microscope images of OSCC tumors for the indicated markers (CK, pan-cytokeratin). KI67$^+$PROX1$^+$ LECs are encircled. Scale bars, 20 μm (upper) and 50 μm (lower).

(C) Kaplan-Meier analyses of RFS, DSS, and OS in patients with T1–T2 OSCC and high or low numbers of juxtatumoral proliferating KI67$^+$PROX1$^+$ LECs (LECp(JT)%). The *p* values are shown for the 3-year time point.

(D) Survival analyses (as in C) of patients stratified by the total number of LECs.

(E) Cox proportional hazard models for 3-year survival in patients with T1–T2 OSCC.

(F) Kaplan-Meier and Cox proportional hazard analyses for 3-year RFS in all patients with OSCC (T1–T4).

(G) Cox proportional hazard analyses for 3-year RFS with a combination risk score (high LECp(JT)% and positive nodal status) and its two components for all patients with OSCC (T1–T4).

(H) Kaplan-Meier analyses of the high-risk group (high LECp(JT)% and positive nodal status), low-risk group (low LECp(JT)% and N0 nodal status), and others (positive only for one of the two risk factors). The *p* values are for the 3-year time point.

(I) Cox proportional hazard analyses of juxtatumoral KI67$^+$PROX1$^+$ LECs in different treatment arms in combined discovery and validation cohorts (T1–T4 patients). HRs represent 3-year RFS (black line) and OS (red line).

(J) Kaplan-Meier analyses of RFS and OS in all patients (T1–T4) with high or low numbers of juxtatumoral KI67$^+$PROX1$^+$ LECs in the surgery-only (upper row) and neck dissection (lower row) treatment groups in the combined discovery and validation cohorts (T1–T4 patients). Kaplan-Meier analyses of RFS are also shown separately for T1–T2 patients.

(K) ROC analysis of the LEC biomarker, TNM stage, T stage, and N status for RFS and OS. The arrow indicates a predetermined cutoff (i.e., LECp(JT)% of 1%). Bar plots show the median and upper and lower quartiles, and the whiskers indicate ±1.5 × IQRs for cumulative 5-year AUC values for RFS (Wilcoxon signed-rank test) in all patients (upper) and in patients with TNM stages I and II in the combined discovery and validation cohorts.

Note the logarithmic HR scale in (E), (F), (G), and (I). See also Figure S7; Tables S4 and S5.

enrichment of the proliferating LECs over nonproliferating LECs at the juxtatumoral location, noted that KI67$^+$PROX1$^+$ LECs were located nearer to the tumor-stroma interface than KI67$^-$PROX1$^+$ cells, and found that LEC proliferation percentages determined from different locations along the interface were concordant (Figures S7I–S7K).

To analyze LEC biomarker performance in different treatment arms, we used combined OSCC data (discovery and validation cohorts pooled, n = 282 patients). Juxtatumoral KI67$^+$PROX1$^+$ LECs had prognostic value for RFS and OS in T1–T4 patients treated with surgery only, surgery combined with radiotherapy, and neck dissection combined with any form of therapy (Figures 6I and 6J). In patients treated with adjuvant chemoradiotherapy, LEC proliferation status was not significantly correlated to survival (Figure 6I). The biomarker also performed well when early-stage T1–T2 patients treated with surgery only or undergoing neck dissection were analyzed separately (Figure 6J). For instance, among T1–T2 patients who underwent neck dissection, only 4% in the KI67$^+$PROX1$^+$ low group experienced disease recurrence in 3 years, whereas 47% of the patients in the KI67$^+$PROX1$^+$ high group experienced relapse (Figure 6J).

Finally, we compared the performance of the KI67$^+$PROX1$^+$ biomarker with that of TNM staging in our pooled OSCC cohort. Receiver operating characteristic (ROC) analyses showed that the LEC biomarker had area under the curve (AUC) values of 0.64 and 0.67 for 3-year RFS and OS, respectively (Figure 6K). The performance of the LEC biomarker and TNM staging was comparable among all patients with OSCC; however, among patients with TNM stages I and II, the LEC biomarker outperformed TNM staging based on cumulative 5-year AUC values (Figure 6K).

Taken together, the number of proliferating LECs at the IM is a robust independent factor predicting survival in early-stage OSCC in both the discovery and validation cohorts. The LEC biomarker can be readily determined by regular immunostaining and is also valuable for predicting the outcomes of patients with more advanced stages of OSCC and in several treatment modalities.

DISCUSSION

The discovery of outcome markers for early-stage OSCC is an unmet clinical need. Although ~20% of patients with early-stage OSCC relapse after primary surgical therapy, no molecular biomarkers exist for patient stratification and therapy development.[2,10] In this study, we provide a multiplexed, spatially resolved atlas of oral cavity cancer TME and report that the number of proliferating LECs (KI67$^+$PROX1$^+$ cells) in the IM of the tumor has strong independent value in predicting survival, especially in early-stage OSCC, in separate discovery and validation cohorts. The LEC biomarker performed well in both non-smokers and smokers, in different treatment modalities, and in comparison with TNM staging. Moreover, we demonstrated that the immunohistological scoring of KI67$^+$PROX1$^+$ LECs is compatible with routine pathological laboratory methodology and is thus a feasible biomarker for clinical decision-making.

Early-stage OSCC is treated with local surgery, and patients typically receive multimodal treatment (neck dissection and adjuvant chemoradiotherapy) only if the tumors are predicted to be aggressive.[2,40] Clinical and pathological TNM staging remain the cornerstones of assessing prognosis and planning and evaluating treatment in OSCC.[2] Additional conventional histopathological patterns like lymphovascular and neural invasion, tumor budding, infiltrative growth pattern, and tumor-to-stromal tissue ratio have been suggested to provide additional prognostic value in early-stage OSCC.[41,42] Out of >150 proposed biomarkers tested in OTSCC, only tumor-cell intrinsic cyclin D1 and VEGFA have emerged as putative biomarkers in meta-analyses, but they are not used clinically.[10]

Multiplexed spatial staining techniques now allow accurate assignment of LECs for the first time. In particular, PROX1, as the master transcription factor for lymphatic vessel development,[28] is highly useful for discriminating LECs from BECs, which otherwise can share many markers (e.g., CD31, LYVE1, PDPN, and VEGFR3).[43,44] We and others previously reported that the total number of lymphatic vessels—defined by commonly used LYVE1 or PDPN markers—had no predictive value in adjusted analyses of OTSCC.[45–48] Consistent with these data, we identified four PROX1$^+$ LEC subsets according to PDPN and LYVE1 expression levels and showed that they have no prognostic value for disease recurrence in early-stage OSCC. Moreover, the total number of LECs had absolutely no association with any of the survival outcomes either in the discovery or validation series.

As such, cell proliferation was not a useful biomarker in our analyses. Specifically, proliferating BECs or tumor cells had no prognostic value in OSCC. This finding agrees with meta-analysis data from 15 studies showing that KI67 (determined in all cells) is not an informative biomarker in mixed stage I–IV OTSCC tumors, or specifically in two smaller cohorts of T1–T2 patients.[10] Moreover, in scRNA-seq and other suspension-based omics analyses of tumors, the number of ECs remains low, and the representativeness of the digested cells for the *in situ* repertoire of LECs remains unverified (e.g., proliferating LECs may be more vulnerable to anoikis during isolation).[30–32] Therefore, spatially resolved single-cell analyses of LECs and precise colocalization of PROX1 and KI67 signals within the cell nucleus have potential benefits in the development of biomarker analyses.

The six tumor cell types and six NESC types (most likely representing fibroblasts) we were able to define using our marker panel did not correlate with OSCC outcome. However, more refined analyses may reveal biomarker values for these cell types. For instance, scRNA-seq[30–32] and multiplexed spatial analyses[49] have suggested partial epithelial-mesenchymal transition as a possible prognostic factor in head and neck cancer (including OSCC); however, the utility of such complex omics-based scores in clinical practice is currently limited. We also found that the regional presence of intratumoral T cells of either the cytotoxic or regulatory phenotype is a good prognostic sign, whereas markers of T cell exhaustion did not confer predictive power. Although counterintuitive, high numbers of Tregs are a good prognostic sign in pan-tumor analysis[50] and OSCC.[51] It is possible that intratumoral Tregs are co-recruited with other T cells by the same signals into the favorable inflamed TME or that they exert beneficial functions by suppressing chronic pro-tumorigenic inflammation.

The prospective nature of our IMC cohort, spatially resolved multiplexed interrogation of the TME, adjusted analyses of clinically challenging early-stage T1–T2 OSCC, and use of an independent validation cohort with cutoff values determined from the discovery cohort are strengths of our analyses. From a clinical point of view, it is highly relevant that we were able to derive an easy-to-measure biomarker from a complex proteomics-based discovery study. Three-color immunofluorescence staining for PROX1, KI67, and tumor cells, low-resolution wide-field microscopy, and manual image analyses were sufficient to replicate the independent prognostic value of proliferating LECs.

In conclusion, here we analyzed the architecture of TME in early-stage OSCC using spatial single-cell technologies. Most notably, we discovered that the KI67$^+$PROX1$^+$ phenotype of LECs at the tumor IM has as an independent value for predicting poor RFS, DSS, and OS. This robust biomarker should be useful in identifying patients with indolent-looking, surgically removed early-stage OSCC who are at risk of progression to fatal disease.

Limitations of the study

There are several limitations to our study. The generalizability of the results to other non-Finnish populations remains to be determined. The low number of different OSCC types (apart from tongue carcinoma) currently prevents meaningful cancer subtype-specific analyses. The value of the LEC biomarker in the constantly changing landscape of OSCC treatments, e.g., in patients treated with immune checkpoint inhibitors, which were not used at the time of cohort collection, also needs to be addressed in future studies. Moreover, the molecular mechanisms that drive LEC proliferation and contribute to decreased survival remain to be elucidated.[52] Finally, the clinical applicability of our two-component biomarker in a pathology laboratory as a stand-alone test in predicting OSCC prognosis after surgery remains to be validated.

RESOURCE AVAILABILITY

Lead contact

Requests for further information, resources, and reagents should be directed to and will be fulfilled by the lead contact, Marko Salmi (marko.salmi@utu.fi).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data supporting the findings of this study are available in the article and its supplementary files. IMC data, including raw segmentation output, segmentation masks, and processed single-cell data, have been deposited in Zenodo (https://doi.org/10.5281/zenodo.14980832). Raw image data are not publicly available since they contain sensitive patient data but are available upon requests provided that data transfer is in agreement with Finnish and EU legislation, informed consent, and institutional review board approvals.
- All original code is uploaded to the GitHub repository (https://github.com/jonasiaho/WholeTumorSeg). The previously established code used in this study has been appropriately referenced in the article.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank Auria Biobank for OSCC samples, Histocore for cutting sections, Turku Bioscience for IMC microscope facility, and CSC for supercomputer time. We thank Dr. Jutta Huvila for providing advice regarding OSCC histopathology. J.N. was financially supported by the TuDMM postgraduate school and the Research Council of Finland InFLAMES Flagship. M.S. obtained research funding from the Finnish Cancer Foundation, Sigrid Juselius Foundation, and InFLAMES Flagship; H.I. and S.V. from Kirsti and Tor Johanssons Hjärt och Cancerstiftelse sr and the State Research Fund for Southwest Welfare District; and S.V. from the Finnish Medical Foundation, Jane and Aatos Erkko Foundation, and the Research Council of Finland (#354599).

AUTHOR CONTRIBUTIONS

Conceptualization, J.N. and M.S.; methodology, J.N.; formal analysis, J.N. and L.N.; investigation, J.N.; resources, S.V., H.I., and M.S.; writing – original draft, J.N. and M.S.; writing – review and editing, all authors; visualization, J.N.; funding acquisition, S.V., H.I., and M.S.; supervision, M.S.

DECLARATION OF INTERESTS

J.N. and M.S. are inventors on a patent application filed by the University of Turku based on this work. Outside this work, S.V. is a co-founder of Thestra, a biotech company focusing on tumor diagnostics, and H.I. is a board member of the company.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Clinical cohorts and sample collection
- METHOD DETAILS
  - Study design
  - Quality control
  - IMC staining and imaging
  - Cell segmentation for IMC
  - IMC data analysis and transformation
  - Clustering of cell phenotypes
  - Cell type identification
  - Cell-cell pairwise interaction analysis
  - Supervised spatial analysis
  - Spatial neighborhoods
  - Defining parameters for validation analyses
  - Immunofluorescence staining
  - Microscopy and quantitative image analysis
  - scRNA-seq analyses
  - Survival analyses
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2026.102615.

Received: June 26, 2025
Revised: November 20, 2025
Accepted: January 13, 2026
Published: February 17, 2026

REFERENCES

1. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., and Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN

estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. *74*, 229–263. https://doi.org/10.3322/caac.21834.

2. Johnson, D.E., Burtness, B., Leemans, C.R., Lui, V.W.Y., Bauman, J.E., and Grandis, J.R. (2020). Head and neck squamous cell carcinoma. Nat. Rev. Dis. Primers *6*, 92. https://doi.org/10.1038/s41572-020-00224-3.
3. GBD 2019 Lip Oral and Pharyngeal Cancer Collaborators; Cunha, A.R.d., Compton, K., Xu, R., Mishra, R., Drangsholt, M.T., Antunes, J.L.F., Kerr, A.R., Acheson, A.R., Lu, D., et al. (2023). The Global, Regional, and National Burden of Adult Lip, Oral, and Pharyngeal Cancer in 204 Countries and Territories: A Systematic Analysis for the Global Burden of Disease Study 2019. JAMA Oncol. *9*, 1401–1416. https://doi.org/10.1001/jamaoncol.2023.2960.
4. Henick, B.S., Taylor, A.M., Nakagawa, H., Wong, K.K., Diehl, J.A., and Rustgi, A.K. (2025). Squamous cell cancers of the aero-upper digestive tract: A unified perspective on biology, genetics, and therapy. Cancer Cell *43*, 178–194. https://doi.org/10.1016/j.ccell.2025.01.003.
5. Rusthoven, K., Ballonoff, A., Raben, D., and Chen, C. (2008). Poor prognosis in patients with stage I and II oral tongue squamous cell carcinoma. Cancer *112*, 345–351. https://doi.org/10.1002/cncr.23183.
6. D'Cruz, A.K., Vaish, R., Kapre, N., Dandekar, M., Gupta, S., Hawaldar, R., Agarwal, J.P., Pantvaidya, G., Chaukar, D., Deshmukh, A., et al. (2015). Elective versus Therapeutic Neck Dissection in Node-Negative Oral Cancer. N. Engl. J. Med. *373*, 521–529. https://doi.org/10.1056/NEJMoa1506007.
7. Amin, MB, e.a.e.A.C.S.M. (2017). American Joint Committee on Cancer and Springer International Publishing, 2017, 8th edn (AJCC Cancer Staging Manual).
8. Moeckelmann, N., Ebrahimi, A., Tou, Y.K., Gupta, R., Low, T.H.H., Ashford, B., Ch'ng, S., Palme, C.E., and Clark, J.R. (2018). Prognostic implications of the 8th edition American Joint Committee on Cancer (AJCC) staging system in oral cavity squamous cell carcinoma. Oral Oncol. *85*, 82–86. https://doi.org/10.1016/j.oraloncology.2018.08.013.
9. Almangush, A., Pirinen, M., Youssef, O., Mäkitie, A.A., and Leivo, I. (2020). Risk stratification in oral squamous cell carcinoma using staging of the eighth American Joint Committee on Cancer: Systematic review and meta-analysis. Head Neck *42*, 3002–3017. https://doi.org/10.1002/hed.26344.
10. Almangush, A., Heikkinen, I., Mäkitie, A.A., Coletta, R.D., Läärä, E., Leivo, I., and Salo, T. (2017). Prognostic biomarkers for oral tongue squamous cell carcinoma: a systematic review and meta-analysis. Br. J. Cancer *117*, 856–866. https://doi.org/10.1038/bjc.2017.244.
11. de Visser, K.E., and Joyce, J.A. (2023). The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. Cancer Cell *41*, 374–403. https://doi.org/10.1016/j.ccell.2023.02.016.
12. Bejarano, L., Jordão, M.J.C., and Joyce, J.A. (2021). Therapeutic Targeting of the Tumor Microenvironment. Cancer Discov. *11*, 933–959. https://doi.org/10.1158/2159-8290.CD-20-1808.
13. Mellman, I., Chen, D.S., Powles, T., and Turley, S.J. (2023). The cancer-immunity cycle: Indication, genotype, and immunotype. Immunity *56*, 2188–2205. https://doi.org/10.1016/j.immuni.2023.09.011.
14. Salmon, H., Remark, R., Gnjatic, S., and Merad, M. (2019). Host tissue determinants of tumour immunity. Nat. Rev. Cancer *19*, 215–227. https://doi.org/10.1038/s41568-019-0125-9.
15. Patras, L., Shaashua, L., Matei, I., and Lyden, D. (2023). Immune determinants of the pre-metastatic niche. Cancer Cell *41*, 546–572. https://doi.org/10.1016/j.ccell.2023.02.018.
16. Caligiuri, G., and Tuveson, D.A. (2023). Activated fibroblasts in cancer: Perspectives and challenges. Cancer Cell *41*, 434–449. https://doi.org/10.1016/j.ccell.2023.02.015.
17. Plikus, M.V., Wang, X., Sinha, S., Forte, E., Thompson, S.M., Herzog, E.L., Driskell, R.R., Rosenthal, N., Biernaskie, J., and Horsley, V. (2021). Fibroblasts: Origins, definitions, and functions in health and disease. Cell *184*, 3852–3872. https://doi.org/10.1016/j.cell.2021.06.024.
18. Oliver, G., Kipnis, J., Randolph, G.J., and Harvey, N.L. (2020). The Lymphatic Vasculature in the 21(st) Century: Novel Functional Roles in Homeostasis and Disease. Cell *182*, 270–296. https://doi.org/10.1016/j.cell.2020.06.039.
19. Zeng, Q., Mousa, M., Nadukkandy, A.S., Franssens, L., Alnaqbi, H., Alshamsi, F.Y., Safar, H.A., and Carmeliet, P. (2023). Understanding tumour endothelial cell heterogeneity and function from single-cell omics. Nat. Rev. Cancer *23*, 544–564. https://doi.org/10.1038/s41568-023-00591-5.
20. Milosevic, V., Edelmann, R.J., Fosse, J.H., Östman, A., and Akslen, L.A. (2022). Molecular phenotypes of endothelial cells in malignant tumors. Biomarkers of the Tumor Microenvironment (Springer International Publishing), pp. 31–52. https://doi.org/10.1007/978-3-030-98950-7_3.
21. Giesen, C., Wang, H.A.O., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., Schüffler, P.J., Grolimund, D., Buhmann, J.M., Brandt, S., et al. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat. Methods *11*, 417–422. https://doi.org/10.1038/nmeth.2869.
22. Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., et al. (2019). ilastik: interactive machine learning for (bio)image analysis. Nat. Methods *16*, 1226–1232. https://doi.org/10.1038/s41592-019-0582-9.
23. Lambert, A.W., Zhang, Y., and Weinberg, R.A. (2024). Cell-intrinsic and microenvironmental determinants of metastatic colonization. Nat. Cell Biol. *26*, 687–697. https://doi.org/10.1038/s41556-024-01409-8.
24. Jackson, H.W., Fischer, J.R., Zanotelli, V.R.T., Ali, H.R., Mechera, R., Soysal, S.D., Moch, H., Muenst, S., Varga, Z., Weber, W.P., and Bodenmiller, B. (2020). The single-cell pathology landscape of breast cancer. Nature *578*, 615–620. https://doi.org/10.1038/s41586-019-1876-x.
25. Schapiro, D., Jackson, H.W., Raghuraman, S., Fischer, J.R., Zanotelli, V.R.T., Schulz, D., Giesen, C., Catena, R., Varga, Z., and Bodenmiller, B. (2017). histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. Nat. Methods *14*, 873–876. https://doi.org/10.1038/nmeth.4391.
26. Patrick, E., Canete, N.P., Iyengar, S.S., Harman, A.N., Sutherland, G.T., and Yang, P. (2023). Spatial analysis for highly multiplexed imaging data to identify tissue microenvironments. Cytometry. A. *103*, 593–599. https://doi.org/10.1002/cyto.a.24729.
27. Giles, J.R., Globig, A.M., Kaech, S.M., and Wherry, E.J. (2023). CD8(+) T cells in the cancer-immunity cycle. Immunity *56*, 2231–2253. https://doi.org/10.1016/j.immuni.2023.09.005.
28. Wigle, J.T., and Oliver, G. (1999). Prox1 function is required for the development of the murine lymphatic system. Cell *98*, 769–778. https://doi.org/10.1016/s0092-8674(00)81511-1.
29. Cousin, N., Cap, S., Dihr, M., Tacconi, C., Detmar, M., and Dieterich, L.C. (2021). Lymphatic PD-L1 Expression Restricts Tumor-Specific CD8(+) T-cell Responses. Cancer Res. *81*, 4133–4144. https://doi.org/10.1158/0008-5472.CAN-21-0633.
30. Puram, S.V., Tirosh, I., Parikh, A.S., Patel, A.P., Yizhak, K., Gillespie, S., Rodman, C., Luo, C.L., Mroz, E.A., Emerick, K.S., et al. (2017). Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. Cell *171*, 1611–1624.e24. https://doi.org/10.1016/j.cell.2017.10.044.
31. Peng, Y., Xiao, L., Rong, H., Ou, Z., Cai, T., Liu, N., Li, B., Zhang, L., Wu, F., Lan, T., et al. (2021). Single-cell profiling of tumor-infiltrating TCF1/TCF7(+) T cells reveals a T lymphocyte subset associated with tertiary lymphoid structures/organs and a superior prognosis in oral cancer. Oral Oncol. *119*, 105348. https://doi.org/10.1016/j.oraloncology.2021.105348.
32. Choi, J.H., Lee, B.S., Jang, J.Y., Lee, Y.S., Kim, H.J., Roh, J., Shin, Y.S., Woo, H.G., and Kim, C.H. (2023). Single-cell transcriptome profiling of the stepwise progression of head and neck cancer. Nat. Commun. *14*, 1055. https://doi.org/10.1038/s41467-023-36691-x.
33. Pencheva, N., Tran, H., Buss, C., Huh, D., Drobnjak, M., Busam, K., and Tavazoie, S.F. (2012). Convergent multi-miRNA targeting of ApoE drives

LRP1/LRP8-dependent melanoma metastasis and angiogenesis. Cell *151*, 1068–1082. https://doi.org/10.1016/j.cell.2012.10.028.

34. Wilson, P.M., Danenberg, P.V., Johnston, P.G., Lenz, H.J., and Ladner, R.D. (2014). Standing the test of time: targeting thymidylate biosynthesis in cancer therapy. Nat. Rev. Clin. Oncol. *11*, 282–298. https://doi.org/10.1038/nrclinonc.2014.51.
35. Liau, N.P.D., Laktyushin, A., Lucet, I.S., Murphy, J.M., Yao, S., Whitlock, E., Callaghan, K., Nicola, N.A., Kershaw, N.J., and Babon, J.J. (2018). The molecular basis of JAK/STAT inhibition by SOCS1. Nat. Commun. *9*, 1558. https://doi.org/10.1038/s41467-018-04013-1.
36. Kwantwi, L.B., Boafo, J.D., Egleh, B.E., and Li, M. (2025). CCL20 in the tumor microenvironment: implications for cancer progression and therapeutic approaches. Clin. Transl. Oncol. *27*, 3285–3292. https://doi.org/10.1007/s12094-025-03874-5.
37. Hippe, A., Braun, S.A., Oláh, P., Gerber, P.A., Schorr, A., Seeliger, S., Holtz, S., Jannasch, K., Pivarcsi, A., Buhren, B., et al. (2020). EGFR/Ras-induced CCL20 production modulates the tumour microenvironment. Br. J. Cancer *123*, 942–954. https://doi.org/10.1038/s41416-020-0943-2.
38. Li, J., Zhang, L., Zhang, Y., Liu, Y., Zhang, H., Wei, L., Shen, T., Jiang, C., and Zhu, D. (2016). A20 deficiency leads to angiogenesis of pulmonary artery endothelial cells through stronger NF-kappaB activation under hypoxia. J. Cell Mol. Med. *20*, 1319–1328. https://doi.org/10.1111/jcmm.12816.
39. Singla, B., Aithbathula, R.V., Pervaiz, N., Kathuria, I., Swanson, M., Ekuban, F.A., Ahn, W., Park, F., Gyamfi, M., Cherian-Shaw, M., et al. (2023). CD47 Activation by Thrombospondin-1 in Lymphatic Endothelial Cells Suppresses Lymphangiogenesis and Promotes Atherosclerosis. Arterioscler. Thromb. Vasc. Biol. *43*, 1234–1250. https://doi.org/10.1161/ATVBAHA.122.318904.
40. Chen, W., Wei, Q., Xiao, T., Lai, J., Huang, M., Ma, Y., Zhang, L., Xue, W., Liu, S., Sun, L., et al. (2025). Evaluating the efficacy and safety of immune checkpoint inhibitors in first and second-line treatments for recurrent and metastatic head and neck squamous cell carcinoma: a systematic review and network meta-analysis of RCTs with a focus on PD-L1 expression. Front. Immunol. *16*, 1508885. https://doi.org/10.3389/fimmu.2025.1508885.
41. Almangush, A., Pirinen, M., Heikkinen, I., Mäkitie, A.A., Salo, T., and Leivo, I. (2018). Tumour budding in oral squamous cell carcinoma: a meta-analysis. Br. J. Cancer *118*, 577–586. https://doi.org/10.1038/bjc.2017.425.
42. Dolens, E.D.S., Dourado, M.R., Almangush, A., Salo, T.A., Gurgel Rocha, C.A., da Silva, S.D., Brennan, P.A., and Coletta, R.D. (2021). The Impact of Histopathological Features on the Prognosis of Oral Squamous Cell Carcinoma: A Comprehensive Review and Meta-Analysis. Front. Oncol. *11*, 784924. https://doi.org/10.3389/fonc.2021.784924.
43. Kalucka, J., de Rooij, L.P.M.H., Goveia, J., Rohlenova, K., Dumas, S.J., Meta, E., Conchinha, N.V., Taverna, F., Teuwen, L.A., Veys, K., et al. (2020). Single-Cell Transcriptome Atlas of Murine Endothelial Cells. Cell *180*, 764–779.e20. https://doi.org/10.1016/j.cell.2020.01.015.
44. Petrova, T.V., and Koh, G.Y. (2020). Biological functions of lymphatic vessels. Science *369*, eaax4063. https://doi.org/10.1126/science.aax4063.
45. Dunkel, J., Vaittinen, S., Grénman, R., Kinnunen, I., and Irjala, H. (2013). Prognostic markers in stage I oral cavity squamous cell carcinoma. Laryngoscope *123*, 2435–2441. https://doi.org/10.1002/lary.23888.
46. de Vicente, J.C., Santamarta, T.R., Rodrigo, J.P., García-Pedrero, J.M., Allonca, E., and Blanco-Lorenzo, V. (2015). Expression of podoplanin in the invasion front of oral squamous cell carcinoma is not prognostic for survival. Virchows Arch. *466*, 549–558. https://doi.org/10.1007/s00428-015-1746-3.
47. Ding, L., Zhang, Z., Shang, D., Cheng, J., Yuan, H., Wu, Y., Song, X., and Jiang, H. (2014). alpha-Smooth muscle actin-positive myofibroblasts, in association with epithelial-mesenchymal transition and lymphogenesis, is a critical prognostic parameter in patients with oral tongue squamous cell carcinoma. J. Oral Pathol. Med. *43*, 335–343. https://doi.org/10.1111/jop.12143.
48. Almahmoudi, R., Kasanen, M., Sieviläinen, M., Salem, A., Pirinen, M., Salo, T., and Al-Samadi, A. (2019). Prognostic value of blood and lymphatic vessel markers in tongue cancer: A systematic review. Cancer Sci. *110*, 3424–3433. https://doi.org/10.1111/cas.14189.
49. Punovuori, K., Bertillot, F., Miroshnikova, Y.A., Binner, M.I., Myllymäki, S.M., Follain, G., Kruse, K., Routila, J., Huusko, T., Pellinen, T., et al. (2024). Multiparameter imaging reveals clinically relevant cancer cell-stroma interaction dynamics in head and neck cancer. Cell *187*, 7267–7284.e20. https://doi.org/10.1016/j.cell.2024.09.046.
50. Gentles, A.J., Newman, A.M., Liu, C.L., Bratman, S.V., Feng, W., Kim, D., Nair, V.S., Xu, Y., Khuong, A., Hoang, C.D., et al. (2015). The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat. Med. *21*, 938–945. https://doi.org/10.1038/nm.3909.
51. Boxberg, M., Leising, L., Steiger, K., Jesinghaus, M., Alkhamas, A., Mielke, M., Pfarr, N., Götz, C., Wolff, K.D., Weichert, W., and Kolk, A. (2019). Composition and Clinical Impact of the Immunologic Tumor Microenvironment in Oral Squamous Cell Carcinoma. J. Immunol. *202*, 278–291. https://doi.org/10.4049/jimmunol.1800242.
52. Ma, Q., Dieterich, L.C., and Detmar, M. (2018). Multiple roles of lymphatic vessels in tumor progression. Curr. Opin. Immunol. *53*, 7–12. https://doi.org/10.1016/j.coi.2018.03.018.
53. Mylly, M., Nissi, L., Huusko, T., Routila, J., Vaittinen, S., Irjala, H., Leivo, I., and Ventelä, S. (2022). Epidemiological Study of p16 Incidence in Head and Neck Squamous Cell Carcinoma 2005-2015 in a Representative Northern European Population. Cancers (Basel) *14*, 5717. https://doi.org/10.3390/cancers14225717.
54. Stirling, D.R., Swain-Bowden, M.J., Lucas, A.M., Carpenter, A.E., Cimini, B.A., and Goodman, A. (2021). CellProfiler 4: improvements in speed, utility and usability. BMC Bioinf. *22*, 433. https://doi.org/10.1186/s12859-021-04344-9.
55. Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. Nat. Methods *18*, 100–106. https://doi.org/10.1038/s41592-020-01018-x.
56. Windhager, J., Zanotelli, V.R.T., Schulz, D., Meyer, L., Daniel, M., Bodenmiller, B., and Eling, N. (2023). An end-to-end workflow for multiplexed image processing and analysis. Nat. Protoc. *18*, 3565–3613. https://doi.org/10.1038/s41596-023-00881-0.
57. Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, E.a.D., Tadmor, M.D., Litvin, O., Fienberg, H.G., Jager, A., Zunder, E.R., et al. (2015). Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. Cell *162*, 184–197. https://doi.org/10.1016/j.cell.2015.05.047.
58. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat. Biotechnol. *42*, 293–304. https://doi.org/10.1038/s41587-023-01767-y.
59. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.
60. Nissi, L., Tuominen, S., Routila, J., Huusko, T., Ketonen, P., Sundvall, M., Leivo, I., Irjala, H., Minn, H., Grönroos, T.J., and Ventelä, S. (2024). xCT as a Predictor for Survival in a Population-Based Cohort of Head and Neck Squamous Cell Carcinoma. Cancer Med. *13*, e70371. https://doi.org/10.1002/cam4.70371.
61. Sauerbrei, W., Taube, S.E., McShane, L.M., Cavenagh, M.M., and Altman, D.G. (2018). Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): An Abridged Explanation and Elaboration. J. Natl. Cancer Inst. *110*, 803–811. https://doi.org/10.1093/jnci/djy088.
62. Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat. Biotechnol. *36*, 421–427. https://doi.org/10.1038/nbt.4091.

63. Schulz, D., Zanotelli, V.R.T., Fischer, J.R., Schapiro, D., Engler, S., Lun, X.K., Jackson, H.W., and Bodenmiller, B. (2018). Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry. Cell Syst. *6*, 25–36.e5. https://doi.org/10.1016/j.cels.2017.12.001.

64. Hoch, T., Schulz, D., Eling, N., Gómez, J.M., Levesque, M.P., and Bodenmiller, B. (2022). Multiplexed imaging mass cytometry of the chemokine milieus in melanoma characterizes features of the response to immunotherapy. Sci. Immunol. *7*, eabk1692. https://doi.org/10.1126/sciimmunol.abk1692.

65. Yang, S., Corbett, S.E., Koga, Y., Wang, Z., Johnson, W.E., Yajima, M., and Campbell, J.D. (2020). Decontamination of ambient RNA in single-cell RNA-seq with DecontX. Genome Biol. *21*, 57. https://doi.org/10.1186/s13059-020-1950-6.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
A full list of IMC antibodies is provided in Table S1	NA	NA
Rabbit monoclonal anti-Prox1 (clone EPR19273)	Abcam	Cat# ab199359; RRID: AB_2868427
Mouse monoclonal anti-Pan Cytokeratin (clone AE1/AE3)	Thermo Fisher Scientific	Cat# 53-9003-82; RRID: AB_1834350
Mouse monoclonal anti-Ki-67 (clone B56)	BD Biosciences	Cat# 558617; RRID: AB_647108
Mouse monoclonal anti-Podoplanin (clone D2-40)	Biolegend	CAT# 916606; RRID: AB_2565820
Rabbit monoclonal anti-TYMS (clone D5B3)	Cell Signaling Technology	Cat# 9045; RRID: AB_2797693
Rabbit monoclonal anti-BIRC5 (clone 71G4B7)	Cell Signaling Technology	Cat# 2808; RRID: AB_2063948
Rabbit monoclonal anti-BST2 (clone E4N3W)	Cell Signaling Technology	Cat# 95940
Donkey anti-Rabbit IgG	Thermo Fisher Scientific	Cat# A31573; RRID: AB_2536183
Goat anti-Rabbit IgG	Thermo Fisher Scientific	Cat# A21039; RRID: AB_2535710
Donkey anti-Mouse IgG	Thermo Fisher Scientific	Cat# A21203; RRID: AB_2535789
Biological samples		
FFPE of 95 early-stage OSCC patients (discovery)	This paper	NA
FFPE of 205 mixed-stage OSCC patients (validation)	Mylly et al.[53]	https://doi.org/10.3390/cancers14225717
Chemicals, peptides, and recombinant proteins		
Dako Target Retrieval Solution (pH 6)	Agilent	Cat# S2369
Dako Target Retrieval Solution (pH 9)	Agilent	Cat# S2367
Iridium intercalator	Standard BioTools	Cat# 201192B
DAPI	Thermo Fisher Scientific	Cat# #D1306
ProLong Gold Antifade Mountant	Thermo Fisher Scientific	Cat# P36930
Critical commercial assays		
Maxpar® X8 Antibody Labeling Kit, 148Nd—4 Rxn	Standard BioTools	Cat# 201148A
Maxpar® X8 Antibody Labeling Kit, 156Gd—4 Rxn	Standard BioTools	Cat# 201156A
Maxpar® X8 Antibody Labeling Kit, 161Nd—4 Rxn	Standard BioTools	Cat# 201161A
Maxpar® X8 Antibody Labeling Kit, 163Nd—4 Rxn	Standard BioTools	Cat# 201163A
Maxpar® X8 Antibody Labeling Kit, 165Nd—4 Rxn	Standard BioTools	Cat# 201165A
Invitrogen Antibody Labeling Kit for 100 µg	Thermo Fisher Scientific	Cat# A88068
Deposited data		
OSCC IMC data	This paper	Zenodo: https://doi.org/10.5281/zenodo.14980832
scRNA-seq data of 18 OSCC samples	Puram et al.[30]	GEO: GSE103322
scRNA-seq data of 6 OSCC samples	Peng et al.[31]	GEO: GSE172577
scRNA-seq data of 10 OSCC samples and 3 control samples	Choi et al.[32]	GEO: GSE181919
Software and algorithms		
R (version 4.3.0)	R Development Core Team	https://www.r-project.org
Ilastik (version 1.3.3)	Berg et al.[22]	https://www.ilastik.org/

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CellProfiler (version 4.2.4)	Stirling et al.[54]	http://cellprofiler.org
Python (version 3.10.9)	Python Software Foundation	https://www.python.org/
Cellpose (version 3.0.11)	Stringer et al.[55]	https://www.cellpose.org/
IMC segmentation pipeline	Windhager et al.[56]	https://github.com/BodenmillerGroup/ImcSegmentationPipeline
imcRtools (version 1.7.2)	Windhager et al.[56]	https://doi.org/10.1038/s41596-023-00881-0
PhenoGraph (version 0.99.1)	Levine et al.[57]	https://doi.org/10.1016/j.cell.2015.05.047
Seurat (version 4.3.0)	Hao et al.[58]	https://doi.org/10.1038/s41587-023-01767-y.
Fiji (version 1.54f)	Schindelin et al.[59]	https://doi.org/10.1038/nmeth.2019
LAS X office (version 1.4.5)	Leica Microsystems	https://www.leica-microsystems.com/products/microscope-software/p/leica-las-x-ls/
Whole tumor segmentation and analysis code	This paper	https://github.com/jonasiaho/WholeTumorSeg

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Clinical cohorts and sample collection

This study included a prospective discovery cohort and a validation cohort. The discovery cohort included all 95 patients with early-stage OSCC who underwent surgical treatment with curative intent in Finland between 2018 and 2019. Samples and clinical data were collected from all university hospitals in Finland, with the latest follow-up updates collected before February 2022. For validation, we used a retrospective cohort of 205 patients with OSCC treated in Finland between 2005 and 2015 at Turku University Hospital. The validation cohort includes all available OSCC samples from a previously described, national, head and neck squamous cell carcinoma (HNSCC) cohort.[53,60] The discovery and validation cohorts are independent with no overlapping patients. All patients were diagnosed by certified hospital pathologists and treated according to the standard of care based on AJCC/UICC staging parameters of OSCC (tumor size, depth of invasion, lymph node status, and extranodal metastasis). Detailed clinicopathological characteristics for both cohorts are presented in Table 1.

All study samples were FFPE whole-tumor sections collected during primary surgery. An OSCC-specialized hospital pathologist guided identification of ROIs from hematoxylin-eosin (H&E)-stained sections for IMC acquisition. The criteria for the selected ROIs were as follows: 1) intact tissue, 2) interface between the invasive tumor front and surrounding stroma (i.e., IM), and 3) lymphocyte-rich area. For the validation cohort, ROI selection was refined using additional immunofluorescence overview images of pan-cytokeratin and PROX1 to include LEC-rich areas at the tumor–stroma interface.

The discovery series was collected with the permission of the Medical Ethics Committee of the Hospital District of Southwest Finland (100/1801/2017). The validation study and utilization of human tissue samples were approved by the Finnish national authority for medicolegal affairs (V39706/2019), regional ethics committee of the University of Turku (51/1803/2017), and Auria Biobank scientific board (AB19-6863). Human lymph node samples were non-tumor cell containing surplus sections from salivary gland and other benign neck surgery (Hospital District of Southwest Finland, permission TO6/033/18). Informed consent was obtained from each participant. All data were handled in a pseudonymized manner, and the image analyses were done in a blinded manner without knowing any clinicopathological data.

METHOD DETAILS

Study design

We analyzed OSCC tumor architecture and searched for prognostic biomarkers in the TME of early-stage OSCC with a focus on immunosuppressive leukocyte types, fibroblasts, and endothelial cells. In our discovery series, we performed spatially resolved single-cell analyses with 25 antibodies using IMC. In an independent validation series, we implemented a conventional immunohistological analysis for the most promising biomarker (KI67$^+$PROX1$^+$ proliferating LECs). We analyzed the value of biomarkers for predicting RFS, DSS, and OS using multivariate models. The study reporting follows the REMARK guidelines[61] (https://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumour-marker-prognostic-studies-remark/; Supplementary Note).

Quality control

All antibodies used in this study underwent an optimization process. We initially tested 41 different antibodies at different concentrations to verify their compatibility with FFPE section staining and to determine optimal signal-to-background ratio. We optimized all antibodies for IMC using OSCC tissue staining. Commercially conjugated antibodies against immune cells and all in-house conjugated antibodies were further optimized using human lymph node tissues with well-defined leukocyte, blood vessel, and lymphatic vessel localization and phenotype patterns (Figures S1A–S1C). The final marker panel was limited to 25 antibodies (Figure 1A) due to costs and to have empty channels between the markers to reduce signal spilling risk. We identified tumor cells using E-cadherin because a commercially available metal-tagged pan-cytokeratin antibody did not work. E-cadherin was expressed in the tumor cells of each sample at a sufficient level to allow cell segmentation. We found inconsistent staining with commercial IMC antibodies for CD276, LAG3, and CD56, and excluded them from the analysis. Antibodies used in the validation cohort for IF were optimized using human lymph node and OSCC tissues.

Four IMC samples in the discovery series were excluded from the analysis due to the absence of tumor cells in the H&E section (n = 2) or irregular staining patterns (n = 2) detected in IMC acquisition. In the validation cohort, 14 samples were excluded due to the absence of tumor cells (n = 5), irregular staining patterns (n = 6), or metastasized disease at the time of diagnosis (n = 3). To ensure robust analyses of proliferating LECs, we also excluded 12 samples that contained minimal or no PROX1$^+$ cells (<20 PROX1$^+$ cells/mm^2 across the entire sample) from the LEC proliferation analysis.

IMC staining and imaging

FFPE sections were deparaffinized in xylene and gradually rehydrated in ethanol, followed by heat-induced epitope retrieval using the Aptum 2100 Retriever and pH 6.0 citrate buffer (#S2369, Dako Target Retrieval). A piece of white paper containing a 6-mm-diameter circle was aligned under the sample slide into the predetermined ROI area from the tissue section (chosen from the IM based on an HE-stained adjacent section). Extra tissue outside the 6-mm circle was gently manually trimmed away using a pipette tip. After gentle washing, the remaining tissue sample was encircled using a hydrophobic ImmEdge PAP pen (#H-4000, Vector Laboratories). We then washed samples with phosphate-buffered saline (PBS) and blocked non-specific binding sites with 5% bovine serum albumin (BSA) for 45 min at room temperature (RT). We prepared a mix of 25 commercial and in-house metal-conjugated antibodies (dilution-range 1:50–1:400, Table S2) in PBS containing 0.5% BSA. We performed all in-house antibody conjugations using BSA- and azide-free commercial antibodies and Maxpar Conjugation Kits (Standard BioTools) according to manufacturer's instructions. We applied the antibody mix on each trimmed section and incubated them overnight at 4°C. After two washings with PBS containing 0.2% Triton X- and two washes in PBS, we incubated the sections with Cell-ID intercalator-Ir (Standard BioTools; 1:200 dilution in PBS) for 30 min at RT. Finally, the slides were washed twice in PBS and stored dry at RT for IMC.

We carried out mass cytometric analyses of the samples with Hyperion imaging system instrument and CyTOF software using standard protocols. We identified the ROIs from optical panorama images generated by Hyperion to verify accurate selection of the acquisition area. IMC data (total 27 channels) was acquired from an area of 1 mm × 1 mm.

Cell segmentation for IMC

We performed data preparation and cell segmentation steps according to IMC Segmentation Pipeline[56] (https://bodenmillergroup.github.io/ImcSegmentationPipeline). Briefly, we created a training set for pixel classification from randomly cropped IMC image stacks of nuclear, cytoplasmic and cell surface channels using Vimentin, CD31, TIM-3, PDPN, E-cadherin, CD8, granzyme B, CD3, IDO1, CD206, CD45, PROX1, and a DNA-binding intercalator marker. Next, we used a trained Ilastik pixel classifier to produce object probability images from full-sized IMC images to distinguish between background, nuclear, and cytoplasmic/cell surface signals and subsequent final cell segmentation masks in CellProfiler (version 4.2.4). We measured single-cell properties and image features in CellProfiler and used them to generate a spatial single-cell dataset in R with imcRtools (version 1.7.2) for downstream analysis.

IMC data analysis and transformation

We performed all IMC data analysis steps in R (version 4.3.0) according to the multiplexed image analysis workflow[56] (https://bodenmillergroup.github.io/IMCDataAnalysis/). All figures were produced in R using ggplot2 (version 3.4.2). We transformed IMC expression values from mean pixel intensity values using an inverse hyperbolic sine function ($y = \sinh^{-1}(x/c)$), where the scaling factor $c = 1$. To remove non-biological batch effects from unsupervised classification of cell phenotypes, we applied batch correction using the mutual nearest neighbor (MNN) method[62] and the mnnCorrect function (batchelor, version 1.16.0) with default parameters.

Clustering of cell phenotypes

We performed unsupervised, marker-based clustering of the cell phenotypes using PhenoGraph (RPhenograph, version 0.99.1), as described previously.[24,57,63] We reduced the batch-corrected expression matrix into principal components (PCA) and used it in PhenoGraph clustering with default parameters and 45 nearest neighbors. Mean marker expressions of the initial clusters were compared and similar phenotypes were aggregated to larger clusters. We performed low-dimensional clustering (three or less markers) with base kmeans function in R.

Cell type identification

We identified the expected major cell populations, including tumor cells (ECAD$^+$), T-cells (CD3$^+$), non-T-cell leukocytes (CD45$^+$CD3$^-$), and LECs (PROX1$^+$) through supervised cell labeling using Cytomapper (version 1.12.0) and training a random forest classifier as described previously[64] using caret package (version 6.0–94) in R. We labeled cells with a maximum classifier probability below 70% as undefined and subsequently clustered them with all markers using PhenoGraph. Based on marker expression and

Cytomapper visualization, we identified two additional distinct major cell populations, CD31$^+$PROX1$^-$ BECs and COL1$^+$ NESCs, and then designated only the remaining cells as undefined cells.

To characterize major cell phenotypes further, we clustered each identified population with PhenoGraph using general and cell type-specific markers. Tumor cells exhibited a distinct PDPN$^+$ subset, which was classified as a separate tumor cell population alongside other PDNP$^-$ ECAD$^+$ tumor cells. Similarly, regulatory T-cells (FoxP3$^+$) and cytotoxic T-cells (CD8A$^+$) formed distinct populations within CD3$^+$ leukocytes, while the remaining CD3$^+$ cells were classified as "T other". In separate representative IMC analyses, approximately 98% of these CD45$^+$CD3$^+$CD8$^-$FoxP3$^-$ "T other" cells expressed CD4 (a marker of T-helper cells), as expected. Overall, we identified nine distinct major cell type lineages (Figure 1B) and 2–5 smaller phenotypic subpopulations for each major cell type (Figure S1A). We used UMAP and t-distributed stochastic neighbor embedding (t-SNE) to visualize the high dimensional marker expression data in two-dimensional space. We computed UMAP and t-SNE dimensions using batch-corrected expression values with Scater package (version 1.28.0).

Cell-cell pairwise interaction analysis

We performed statistical testing of significant pairwise interactions between cell types as described previously[24,25] using testInteractions function (imcRtools, version 1.7.2). In brief, averaged interaction counts between two different cell types was compared against an empirical null distribution generated by cell type permutations. We considered cells within a default radius of 8 pixels (8 μm) interacting. The observed counts were compared to the randomized distribution to determine whether the cell types were found together (attraction) or apart (avoidance) more often than expected by chance. We then defined significant interaction/attraction (maximum score +1) or avoidance (minimum score -1) with p-value <0.01 for each cell type-specific interaction per sample.

Supervised spatial analysis

We identified sets of connected tumor cells as individual tumor patches[63] with the patchDetection function and Steinbock interaction graph (imcRtools, version 1.7.2). Cells within a default 8-pixel radius (8 μm) were considered interacting, and tumor patches containing fewer than 10 cells were excluded. We calculated distance to nearest tumor patch (DTNT) for each cell. CD8$^+$ T-cells with DTNT <0 were labeled as tumor-infiltrating cytotoxic T-cells (Tc IT) and FoxP3$^+$ T-cells with DTNT <0 as tumor-infiltrating regulatory T-cells (Treg IT) in subpopulation-level analyses. We also used the patch-detection approach to detect sets of connected LECs as individual lymphatic vessels.

Spatial neighborhoods

For unsupervised detection and classification of cellular neighborhoods, we implemented a previously described method[26] in which local indicators of spatial association (LISA) are used to cluster cells into similar spatial communities. We did this using the lisaClust function (lisaClust, version 1.8.1) with number of centers (k) set to five and LISA curve radii of 10, 20, 50, and 100 μm. We labeled the tumor cell-enriched spatial neighborhoods as "IM Tumor/central" and "IM Tumor/border," according to their localization relative to the tumor–stroma interface. We classified non-tumor cell-enriched spatial neighborhood clusters based on the most enriched cell types and labeled them as "IM T cell hotspot", "IM Pan-leukocyte hotspot", or "IM Stroma-rich region" (Figure 2C). Endothelial cells (BECs and LECs) and NESCs were the enriched individual cell types in the "IM Stroma-rich region".

Defining parameters for validation analyses

We determined the optimal cut-off values for high and low proliferation using ROC analysis with survivalROC (version 1.0.3.1) for all PROX1$^+$ LECs (LECp%) and juxtatumoral PROX1$^+$ LECs [LECp(JT)%] based on data from the discovery cohort. A distance of approximately 50 μm from the nearest tumor cell was used to define juxtatumoral LECs. We selected the cut-off point that best predicted OSCC recurrence based on Youden's J statistic. To assess the impact of LEC density on the prognostic power of LEC proliferation, we evaluated the AUC for LEC proliferation across multiple density cut-offs.

Immunofluorescence staining

To evaluate the KI67$^+$PROX1$^+$ biomarker in the validation series, FFPE sections underwent deparaffinization and epitope retrieval following the same protocol as used with the IMC samples. We used a buffer containing 2% fetal calf serum (FCS) and 1% BSA in PBS for blocking after antigen retrieval and for preparing antibody dilutions. The sections were first incubated with unconjugated anti-PROX1 antibody (EPR19273, #ab199359 Abcam; 1:500 dilution) overnight at 4°C, washed in PBS, and incubated with Alexa Fluor 647-labeled secondary antibody (#A31573, Invitrogen, 2 μg/mL) for 1 h at RT. After washings, we added Alexa Fluor 488-conjugated anti-pan cytokeratin (AE1/AE3, #53-9003-82 Invitrogen; 1:100 dilution) and Alexa Fluor 555-conjugated anti-KI67 (B56, #558617 BD Pharmingen; 1:50 dilution) antibodies for 1 h at RT. Finally, samples were washed in PBS, and DAPI (#D1306, Invitrogen; 1:5,000 dilution in PBS) was incubated with the sections for 10 min at RT before a wash with deionized water. We mounted the sections with ProLong Gold Antifade Mountant (#P36930, Invitrogen). We stored the samples in dark at 4°C and imaged them within 1 week.

To validate the molecular phenotypes of LECs inferred from the scRNA-seq analyses at the protein level, we immunostained FFPE sections of OSCC tumors using the protocol described above with the following modifications. Antigen retrieval was performed using pH 9.0 Tris/EDTA buffer (#S2367, Dako Target Retrieval). Predetermined optimal dilutions of anti-TYMS (#9045, CST, 1:200 dilution), anti-BIRC5 (#2808, CST, 1:200 dilution), and anti-BST2 (#95940, CST, 1:200 dilution) were used as primary antibodies (one antibody per section). They were visualized using Alexa Fluor 750-conjugated second-stage anti-rabbit IgG antibody (#21039, Thermo Fisher, 1:1,000 dilution). All sections were then stained with DAPI, mouse anti-PDPN (#916606, Biolegend, 1:1,000 dilution), Alexa Fluor 594 anti-mouse IgG antibody (#A21203, Thermo Fisher, 1:1,000 dilution), Alexa Fluor 555-conjugated anti-KI67 antibody (#558617, BD Pharmingen, 1:50 dilution), Alexa Fluor 488-conjugated anti-pan-cytokeratin antibody (#53-9003-82; Thermo Fisher, 1:100 dilution), and in-house conjugated Alexa Fluor 647-conjugated anti-PROX1 antibody (#ab199359, Abcam, 1:200 dilution).

Microscopy and quantitative image analysis

We acquired all image data used for scoring of the validation samples with Leica Thunder wide-field microscope using a 20× Plan Apo 0.80 objective and LAS X imaging software. We used opto-digital computational clearing algorithm (Thunder), included in the LAS X software, to remove hazing artifacts. We down-scaled two 1-mm^2 ROIs best matching the selection criteria by 0.5× and imported them to ImageJ (version 1.54f). We annotated all PROX1$^+$ cells and PROX1$^+$KI67$^+$ cells manually using ImageJ built-in cell counter. We considered all PROX1$^+$KI67$^+$ cells within an approximated 50 μm radius from the nearest tumor cell (i.e., juxtatumoral location). Cell type annotations, along with x-y coordinates of each annotated cell, were saved in.xml files that were used to collect absolute cell numbers for survival analyses.

For analyzing whole tumor sections, we used randomly cropped stacks of IF-stained images from the validation series to train the Cellpose (version 3.0.11) deep learning network to generate cell segmentation masks of 20 whole tumor sections, which were randomly selected from samples with high LEC proliferation in the manually scored data. We performed image processing in Python (version 3.10.9) using numpy (version 1.24.2), tiffile (version 2024.9.20), and pandas (version 1.5.3). For computational efficiency, we downscaled the image resolution by 0.5× and divided whole tumor sections into small tiles, each assigned a grid value to retain original x-y positional information. We measured mean pixel intensities of PROX1, KI67, DAPI, and pan-cytokeratin in CellProfiler (version 4.2.4) along with spatial coordinates. We used the segmented data to generate a spatial single-cell dataset in R (version 4.3.0) and transformed tile-specific spatial coordinates back to whole-tumor scale using the assigned grid values. We assigned uniform marker cutoff values for cell type identification experimentally using Cytomapper (version 1.12.0), and we validated the identification of LEC and LEC KI67$^+$ by comparing manually scored and segmented data (Figure S7G). We defined the juxtatumoral area from pan-cytokeratin images with a binary line drawn to separate tumor and stromal compartments. We skeletonized the binary line using skimage (version 0.24.0) into one-pixel-wide coordinates and used it as a reference for measuring cell distances.

To assess the variation in LEC proliferation at IM, we analyzed LEC proliferation in up to ten 1 mm^2 regions per sample. We then calculated the mean proliferation value for each whole tumor section and determined the mean absolute deviation (LEC proliferation % in one sample area/LEC proliferation % in the whole sample ×100).

For TYMS, BIRC5, and BST2 quantification, LECs were segmented in tumor-containing IM areas and in distant normal regions of OSCC samples using PROX1 and PDPN signals to train the Ilastik pixel classifier, followed by further segmentation in CellProfiler. Samples from 3 to 5 patients (6–8 areas per sample) containing 5,569 LECs were used to quantify the expression of each new marker. The mean pixel intensity within individual cell masks was used to calculate marker expression levels and normalized to the 99th percentile.

We acquired representative high-resolution confocal images using Leica Stellaris 8 Falcon FLIM microscope and HC PL APO 20×/0.75 objective.

scRNA-seq analyses

We reanalyzed publicly available OSCC scRNA-seq datasets (GSE103322, GSE172577, and GSE181919) from the Gene Expression Omnibus (GEO). We created new Seurat (version 4.3.0) objects in R (version 4.3.0) and applied the same quality control criteria for cell viability as reported in the original publications.[30–32] We used the standard Seurat workflow and nearest-neighbor graph-based method to cluster all cells. We identified clusters expressing pan-endothelial markers, including *PECAM1*, *CLDN5*, and *CDH5*. We then isolated endothelial cells and re-clustered them, identifying a PDPN$^+$PROX1$^+$ LEC cluster and a PDPN$^-$ PROX1$^-$ BEC cluster. We classified LECs with an *MKI67* RNA count >0 as KI67$^+$.

We removed ambient RNA contamination from all datasets using the decontX (version 1.4.1) pipeline with default parameters.[65] We identified differentially expressed genes (DEGs) between cell populations using FindMarkers (Seurat v4.3.0) with a minimum expression fraction of 0.20 and a log fold-change threshold of 0.25. We performed Gene Ontology over-representation analysis using clusterProfiler (version 4.14.4) for DEGs with an adjusted p-value <0.05. We analyzed upregulated and downregulated DEGs separately, and considered pathways with a q-value <0.05 to be significantly enriched.

Survival analyses

We define OS as the time from the end of primary treatment to the date of death from any cause or the last follow-up, RFS as the time from the end of primary treatment to the date of disease recurrence or the last follow-up, and DSS as the time from the end of primary treatment to the date of cancer-related death or the last follow-up. We analyzed survival using log rank (Mantel–Cox) test. In the discovery series, we categorized the continuous variables into high and low groups by the median value for Kaplan–Meier analyses. We calculated HR for univariate and multivariate analyses using Cox proportional hazards model.

QUANTIFICATION AND STATISTICAL ANALYSIS

We performed statistical analyses in R (version 4.3.0) and SPSS 27 (SPSS, IBM, Armonk, NY, USA). P-values <0.05 were considered statistically significant. Normal distribution was evaluated using Shapiro-Wilk test. We analyzed parametric data using Student's *t* test and non-parametric data by Mann–Whitney U test for unpaired analyses and Wilcoxon signed rank test for paired analyses. The *t* test was used for large sample sizes when no extreme deviation from a normal distribution was observed. We analyzed contingency tables by Fisher's exact test. We used Pearson correlation to test linear correlation of continuous variables. Statistical analyses from the validation cohort were initially conducted by an independent researcher (L.N.) using fixed parameters and pre-determined cut-off values based on the results from the discovery cohorts. The data from the discovery and validation cohorts were combined in the analysis of different treatment subgroups and LEC biomarker performance against TNM staging. In the ROC and AUC analyses, the LEC biomarker was treated as a continuous variable and the TNM stage as a noncontinuous four-stage variable. The T-class (T1-T4) and N-status (N0-N2) were used as four-stage and three-stage noncontinuous variables, respectively.