The Tumor Immune Microenvironment Architecture Correlates with Risk of Recurrence in Head and Neck Squamous Cell Carcinoma

Donald T. Weed[1,2], Serena Zilio[3], Christie McGee[1], Boutheina Marnissi[3], Zoukaa Sargi[1,2], Elizabeth Franzmann[1,2], Giovana Thomas[1,2,†], Jason Leibowitz[1,2], Elizabeth Nicolli[1,2], David Arnold[1,2], Silvio Bicciato[4], and Paolo Serafini[1,2,3]

ABSTRACT

Emerging evidence suggests that not only the frequency and composition of tumor-infiltrating leukocytes but also their spatial organization might be a major determinant of tumor progression and response to therapy. Therefore, mapping and analyzing the fine tumor immune architecture could potentially provide insights for predicting cancer prognosis. Here, we performed an explorative, prospective clinical study to assess whether structures within the tumor microenvironment can predict recurrence after salvage surgery in head and neck squamous cell carcinoma (HNSCC). The major immune subsets were measured using flow cytometry and co-detection by indexing (CODEX) multiparametric imaging. Flow cytometry underestimated the number of PMN-MDSCs and neutrophils in the tumor and overestimated the tumor-infiltrating lymphocyte frequency. An *ad hoc* computational framework was used to identify and analyze discrete cellular neighborhoods. A high frequency of tertiary lymphoid structures composed of CD31^{high}CD38^{high} plasma cells was associated with reduced recurrence after surgery in HNSCC. These data support the notion that the structural architecture of the tumor immune microenvironment plays an essential role in tumor progression and indicates that type 1 tertiary lymphoid structures and long-lived CD31^{high}CD38^{high} plasma cells are associated with good prognosis in HNSCC.

Significance: Imaging the spatial tumor immune microenvironment and evaluating the presence of type 1 tertiary lymphoid structures enables prediction of recurrence after surgery in patients with head and neck squamous cell carcinoma.

Introduction

The tumor immune microenvironment plays an important role in disease progression and response to therapy (1). For example, a higher concentration of regulatory T cells (Treg; ref. 2) and CD33$^+$IL4α$^+$ myeloid-derived suppressor cells (MDSC) correlate with the risk of recurrence. In contrast, a higher number of CD8$^+$ tumor-infiltrating lymphocytes (TIL) increases recurrence-free survival in head and neck squamous cell carcinoma (HNSCC), breast cancer, lung cancer, and other human malignancies (3–7). Clinical staining of leukocytes is routinely performed with conventional one- or two-color IHC. However, studying more than two markers requires a careful selection of primary antibodies or the use of consecutive tissue sections. This is problematic for studying samples with low tissue availability and makes it extremely difficult to colocalize markers at the single-cell level. Alternatively, the tumor microenvironment is assessed by flow cytometry or single-cell (sc) RNA sequencing on single-cell suspensions from the tumors. These techniques allow for the simultaneous analysis of multiple markers and genes but require tissue dissociation. Thus, they do not reveal the topographic location of the leukocyte subsets and raise concerns about the possible underrepresentation of some leukocyte subsets that may be lost during tissue processing (8, 9). The emergence of multiplexed tissue imaging overcame these challenges. In particular, co-detection by indexing (CODEX) generates detailed information on the distribution of different cellular phenotypes while maintaining the morphologic context of healthy and diseased tissues (10–12). Although this technology has been used primarily on tumor tissues microarrays (11–15) that may exclude functional structures distant from the neoplastic edge, it revealed a complex organization of the tumor microenvironment into spatially defined cellular structures whose composition and frequency could be used to discriminate patients with HNSCC with lymph nodal metastases from those with N0 tumors (10–15).

Here we performed a prospective clinical study of patients with stage II to IV surgically resectable primary or recurrent HNSCC to (i) characterize the tumor microenvironment by multicolor flow cytometry and CODEX on large HNSCC sections; (ii) identify structures distant from the tumor edge; and (iii) evaluate whether the tumor architecture can be associated with prognosis. Compared with CODEX, we found that flow cytometry underestimated the number of neutrophils and polymorphonuclear (PMN) MDSCs and overestimated TILs frequency. Using CODEX data, we identified 20 cell types (CT) and 11 cellular neighborhoods (CN) that portray the HNSCC microenvironment. Moreover, we provided evidence that a

[1]Department of Otolaryngology, Miller School of Medicine, University of Miami, Miami, Florida. [2]Sylvester Comprehensive Cancer Center, Miller School of Medicine, University of Miami, Miami, Florida. [3]Department of Microbiology and Immunology, Miller School of Medicine, University of Miami, Miami, Florida. [4]Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy.

S. Bicciato and P. Serafini contributed equally to this article.

†Deceased.

Corresponding Authors: Paolo Serafini, Microbiology and immunology, University of Miami, 1600 NW 10 Avenue, Miami, FL 33136. E-mail: pserafini@miami.edu; and Silvio Bicciato, Department of Life Sciences, University of Modena and Reggio Emilia, Via G. Campi 287-41125 Modena, Italy. E-mail: silvio.bicciato@unimore.it

"type 1" tertiary lymphoid structure (TLS1) and CD38$^+$CD31$^+$plasma cells are associated with fewer recurrences after salvage surgery in HNSCC.

Materials and Methods

Patients and ethical statement

Nine patients (median age 71, range 61–78) with stage II to IV HNSCC undergoing surgical resection were enrolled in the study under the University of Miami Institutional Review Board (IRB)–approved protocol no. 20200268 after signing an IRB-approved written informed consent. Seven patients had HNSCC in the oral cavity, one in the hypopharynx, and one in the oropharynx. Two patients were female, 5 had a smoking history, 6 had a recurrent tumor, and 5 had prior radiation to the field. No tumor was HPV related. Complete clinical and pathologic parameters are reported in **Table 1**.

Murine tumor model and ethical statements

Eight weeks old Balb/c female mice (Taconic) were injected ortho-topically with 10^5 4T1 (ATCC) cells in the fat pad of the right abdominal mammary gland. Mice were euthanized when tumors reached 0.7 cm in diameter. All experiments with mice were performed under the Institutional Animal Care and Use Committee (IACUC) protocol no. 22051, approved by the ethical committee of the University of Miami. Part of each tumor and spleen was processed for flow cytometry, and part was snap frozen in OCT for subsequent CODEX analysis.

Tissue processing

Human tumors were harvested at surgery and processed within 1 hour. Mouse tissues were processed immediately after euthanasia. Part of the tumor was flash-frozen in OCT compound for subsequent CODEX staining, and part was mechanically dissociated into single-cell suspension for multicolor flow cytometry analysis. Briefly, tumor specimens were washed twice with PBS and incubated for 20 minutes at 37°C with five volumes of PBS containing *Clostridium histolyticum* collagenase type IV (10 mg/mL; Sigma), MgCl$_2$ (100 μmol/L), and CaCl$_2$ (100 μmol/L). The solution was passed every 10′ in a needleless 5 mL syringe. Cells were then passed through a 70 μm cell strainer, washed with PBS, and resuspended for staining.

Flow cytometry

Single-cell suspensions (10^6 cells) from HNSCC tumors were stained for 20′ at 4°C with the Zombie Violet Fixable Viability Dye (BioLegend), washed with PBS, and stained with the following antibody panels for 15′ at 4°C. (i) Myeloid cell panel: Anti-human CD33-FITC (clone HIM3–4; BD), anti-human Lox1-APC (clone 15C4; BioLegend), anti-human CD124-PE (clone 25463; R&D Systems), anti-human CD14-APC-H7 (clone MφP9; BD Biosciences), anti-human CD15-BV711 (clone W6D3; BD Biosciences), anti-human HLA-DR V500 (clone G46–6; BD Biosciences), and anti-human CD11b-BV605 (clone ICRF44; BD Biosciences). (ii) Lymphoid cell panel: Anti-human CD3-Alexa Fluor 700 (clone OKT3; eBioscience), anti-human CD247-PE (clone 6B10.2; eBioscience), anti-human CD4-BV711 (clone SK3; BD Biosciences), anti-human CD8-BV605 (clone SK1; BD Biosciences), anti-human CD69-APC-Cy7 (clone FN50; BD Biosciences), and anti-human CD154-PE/Dazzle 594 (clone24–31; BioLegend). Cells in the lymphoid panel were then fixed and permeabilized using the Foxp3/Transcription Factor Staining Buffer Set (eBioscience) as per the manufacturer's instructions and with the anti-human Foxp3-APC (clone 236A/E7; eBioscience). Mouse tumors and spleens were stained with NIR live/dead dye (Thermo Fisher Scientific)

Table 1. Clinical and pathologic parameters of the patients enrolled in the study.

ID	Site	Pathologic stage AJCC 8	Previous HNSCC tumor	Largest tumor size (cm)	PNI	LVI	DOI (cm)	Previous radiation	Age (years)	Sex	Smoking history	HPV-related	New recurrence	New recurrence (days from surgery)	Follow up (days)
s2	Hypopharynx	pT2N0Mn/a	1	1.1	0	0	0.2	1	62	M	0	No	0		448
s3	Oral cavity (hard palate)	pT3NXMn/a, cT3N0M0	1	2.3	0	0	0.6	1	75	F	1	No	0		741
s4	Oral cavity (tongue)	pT3N2bMn/a	0	3	1	1	1.8	0	77	M	1	No			267
s6	Oral cavity (tongue)	pT3N0M0	0	3.9	1	0	1.7	0	61	F	1	No	1	131	755
s7	Oral cavity (tongue, FOM)	pT2N0M0	1	3	0	0	0.6	1	64	M	0	No	0		
s10	Oropharyngeal	pT1N0Mn/a	1	1.5	1	0		1	64	M	1	No	1	255	410
s11	Oral cavity (maxillary alveolar ridge)	pT3N2bMn/a	0	3	0	0	1.1	0	78	M	0	No	0		725
s12	Oral cavity	pT3NxMn/a	1	5.8	0	0	0.5	0	72	M	1	No	1	338	436
s13	Oral cavity (tongue)	rpT3N0Mn/a	1	2.4	1	0	1.3	1	71	M	0	No	1	222	548

Note: Yes = 1; No = 0; Blank, unknown/not applicable.

and with rat anti-mouse antibodies specific for CD45 (AF700, clone AB_2572116; BioLegend), CD11b (bv711, clone AB_394002; BD Biosciences), Ly6 g (Pac.Blue, clone AB_1727562; BD Biosciences), Ly6c (FITC, clone AB_394628; BD Biosciences), CD3 (PE clone AB_1834427; Thermo Fisher Scientific), CD4 (FITC, clone AB_394585; BD Biosciences), and CD8 (Pac.Blue, clone AB_394571; BD biosciences). Data were acquired within 2 hours on a BD LSRII flow cytometer analyzer equipped with the following wavelengths lasers: 405 nm (50 mW), 488 nm (50 mW), 532 nm (150 mW), and 640 nm (40 mW) or with the 5 lasers Cytoflex LX-19C (Beckman Coulter). Analyses were performed in FCS Express 7 (De Novo software).

CODEX and custom antibodies

The following oligonucleotides conjugated antibodies against human epitopes were acquired from Akoya: CD45-BX001 (HI30), CD19-BX003 (HIB19), CD8-BX004 (SK1), CD38-BX007 (HB-7), CD279-BX014, (EH12.2H7), CD3-BX015 (UCHT1), pan-cytokeratin-BX019 (AE-1/AE-3), CD4-BX021 (SK3), CD90-BX022 (5E10), HLA-DR-BX026 (L243), CD31-BX032 (WM59), CD34-BX035 (561), and CD69-BX041 (FN50). In addition, we generated the following oligonucleotides-tagged antibodies against human epitopes using the CODEX Conjugation Kit (Akoya) as per manufacturer instructions: FOXP3-BX002 (AB_467556; Thermo Fisher Scientific), PDL1-BX005 (AB_2687808; BioXcell), CD14-BX006 (AB_830675; BioLegend), CD163-BX010 (AB_1088991; BioLegend), CD11b-BX013 (AB_314154; BioLegend), CD33-BX016 (AB_2562818; BioLegend), CD56-BX017 (AB_314444; BioLegend), IL4Ra-BX020 (AB_2126871; R&D Systems), CD45RA-BX024 (AB_314406; BioLegend), CD123-BX027 (RRID: AB_467453; Thermo Fisher Scientific), CCR7-BX028 (AB_2563726; BioLegend), LOX1-BX029 (AB_2562272; BioLegend), CD25-BX030 (AB_2561752; BioLegend), CD15-BX033 (AB_314194; BioLegend), CD66b-BX036 (AB_2728422; BioLegend), CD68-BX037 (AB_1089058; BioLegend), and CD16-BX042 (AB_314202; BioLegend). The following antibodies against mouse epitopes were purchased from Akoya: CD45-BX007 (clone AKYP0005), CD3-BX021 (clone AKYP0035), CD4-BX026 (clone AKYP0041), CD8a-BX029 (clone 53–6.7), CD11b-BX025 (cloneM1/17), and Ly6g-BX024 (clone AKYP0039). The anti-mouse Ly6c antibody (clone HK1.4; BioLegend) was conjugated with the BX027 barcode. Conjugation of anti-mouse and anti-human antibodies was verified by PAGE, and optimal titration was performed by single antibody staining on HNSCC tumor specimens of mouse spleen.

CODEX staining, image acquisition, and segmentation

OCT embedded flash-frozen specimens were cut with a cryostat microtome in 5 μm sections and placed on a poly-L-lysine-coated coverslip. Sections were dried with drierite beads for 2′, fixed in acetone for 10′ at RT, rehydrated in hydration buffer (3 × 2′; Akoya), and fixed in prestain fix buffer (1.6% PFA in hydration buffer) for 10′. Sections were washed in hydration buffer and incubated for 20′ in staining buffer (Akoya). Tissues were then stained in a humidity chamber with the cocktail of oligonucleotides tagged antibodies diluted in blocking buffer (N, G, J, S, Akoya) for 3 hours at 4°C. Sections were then washed three times for 2′ in staining buffer and fixed by one incubation poststain fixative buffer (1.6% PFA in staining buffer) for 10′, three washes with PBS, one incubation in ice-cold methanol for 5′, three washed in PBS, and a 20′ incubation with the final fixative solution (Akoya) at RT. Sections were washed nine times in PBS and kept in storage buffer until image acquisition.

Fluorochrome conjugated reported oligonucleotides were hybridized on 12-cycle experiments with the Akoya PhenoCycler connected to a BZ-X700 microscope (Keyence), following manufacturer instructions. Images were automatically acquired on 121 adjacent 20× fields (5.6 mm × 4.2 mm, resolution 377.47 nm/pixel) with 7 z-slices (1.5 μm/z pitch) and 30% tile overlap, with 10, 250, 350, and 350 milliseconds acquisition time for DAPI, AF488, Cy3, and cy5 filters, respectively. The resulting 40,656 images per specimen were processed with Akoya's PhenoCycler Instrument Manager (version 1.3) and CODEX Processor (version 1.7.2) for segmentation using the default parameters.

Composite images were loaded in ImageJ (version 1.53) using the CODEX MAV extension (version 1.5.0.8) to check image quality and export single-cell fluorescence signals as CSV files.

Computational analysis of CODEX data
Data preprocessing and cell phenotyping

Data have been preprocessed using Seurat 4.2.1 functions to analyze image-based spatial data (https://satijalab.org/seurat/articles/spatial_vignette_2.html) in R 4.2.2. Briefly, CSV data files from CODEX MAV have been loaded in R using Seurat *LoadAkoya* function, and protein signals have been normalized with the centered log-ratio based normalization. To detect the cell phenotypes in the HNSCC tissue (i.e., malignant, stromal, immune, or vascular cells), we first performed principal component analysis (PCA) for dimensionality reduction and then cluster analysis in the low-dimensional space. Before applying PCA, we scaled the data and determined the number of components for downstream analysis. Then, we ran dimensional reduction for each sample using the top 11 principal components and graph-based clustering at different resolutions (ranging from 0.2 to 0.8; Supplementary Table S1). Cell clusters have been visualized on a protein intensity-based uniform manifold approximation and projection (UMAP) and on their spatial location. To associate cell phenotypes to clusters, levels of protein markers have been displayed as dot plots for each cluster at any explored clustering resolution. For each sample, the optimal number of clusters (i.e., the clustering resolution) was determined based on visual inspection of cluster location and marker expression (Supplementary Table S1). Clusters with a similar morphological appearance in the tissue and similar marker expression profiles were merged, and artifacts were removed. Next, we manually assigned CT annotations to clusters at the selected resolution based on the average expression of protein markers in each cluster. After removing cells marked as artifacts, raw count and normalized data matrices of all samples have been merged using the *merge* function of the Seurat package. Supervised annotation and merging resulted in a final list of 20 CTs used to annotate all cells in every sample (Supplementary Table S1).

Neighborhood identification

To identify CNs that are regions with a characteristic local composition of cell phenotypes, we implemented in R the neighborhood analysis presented in ref. 11. For each of the 792,349 cells from the nine samples, we identified a window consisting of W nearest neighboring cells (including the center cell) using the *nn2* function of the *RANN* R package (version 2.6.1). The *nn2* function uses a k-dimensional tree to find a given number of near neighbors (here, W) for each point identified by the X and Y coordinates in the input dataset. These windows were then clustered by their composition with respect to the 20 CTs previously identified by graph-based clustering and supervised annotation. Specifically, each window was converted to a vector of length 20 containing the frequency of each of the 20 CTs among the W neighbors. Subsequently, windows have been clustered using the *MiniBatchKmeans* function of the *ClusterR* package (version 1.2.9)

implementing the Mini-batch K-means clustering algorithm with a given value of K. Each cell was then allocated to the CN of its surrounding window using the *predict_MBatchKMeans* function of the *ClusterR* package. We applied the entire procedure and identified CNs for different combinations of W (ranging from 7 to 200) and K (ranging from 7 to 14).

Quantification of pairwise cell–cell contacts

To quantify pairwise cell–cell contacts within CNs, we first determined the direct neighbors of each cell within each CN using the Delaunay triangulation implemented in the *triangulate* function of the *RTriangle* R package (version 1.6–0.11). First, we calculated the number of contacts between cells of types i and j as the set of edges N_{ijk} between cells of type i and j in CN k returned by the triangulation. Then, displayed pairwise contacts among different phenotype cells as circle plots using the *chordDiagram* function of the *circlize* R package (version 0.4.15).

To evaluate if cell–cell contacts of cells assigned to a given phenotype were influenced by the CN, we compared the relative cell–cell contact frequencies (RF) for all phenotypes in each pair of CNs (e.g., TLS1 and TLS2) across all patients containing both CNs. We calculated the relative cell–cell contact frequency between cells of types i and j in each CN k as N_{ijk}/N_{ik} where N_{ijk} is the number of edges of the triangulation between cells of type I and j and N_{ik} is the number of edges of type i in CN k. We compared the relative cell–cell contact frequency in pairs of CNs using the *ttest* function of the R *stats* package. We discarded cell–cell interactions between pairs of phenotypes present in less than three samples. P values have been adjusted for multiple comparisons with the *p.adjust* function of the R *stats* package using the method of Benjamini and Hochberg (16).

Tensor decomposition analysis

We used the procedure described in ref. 14 for tensor decomposition analysis, adapting the Python code available at https://github.com/nolanlab/NeighborhoodCoordination. For each patient, we constructed a joint composition matrix quantifying the frequency of each CT in each CN and collected these matrices in tridimensional tensors with dimensions given by patients by CTs by CNs. We generated distinct tensors for groups of patients with different clinical characteristics. Non-negative Tucker decomposition, as implemented in the *Tensorly* Python package, was applied to each tensor (17). We selected the ranks in each dimension equal to 3, 8, and 8 by visually inspecting the decomposition error for different combinations of the ranks in multiple decomposition runs. Modules for CT and CNs have been generated from the factor matrices of the CT and CN spaces, respectively. The interactions composing a tissue module were obtained from the 8×8 slices of the $3 \times 8 \times 8$ core tensor linking the factor matrices. In the graphical representation of tensor decomposition analysis, the contribution of CTs and CNs to tissue modules was quantified by their values in the corresponding slice of the core tensor, and these values were used for color shading. Elements with a value lower than 0.1 in the corresponding slice of the core tensor have been greyed out.

Differential marker expression

The analysis of differential marker expression has been performed using the *FindMarkers* function of *Seurat* R package (version 4.2.1). To identify differentially expressed markers in the comparison of a given phenotype (e.g., plasma cells) in a CN (e.g., TLS1) versus the same phenotype in all other CNs, we set the fold change and FDR thresholds at 1.2 and 0.05, respectively.

Analysis of gene expression profiles in patients with HNSCC

Raw data for The Cancer Genome Atlas (TCGA) HNSC collection were downloaded as raw counts from the TCGA repository using the *TCGAbiolinks* R package (v. 2.29; ref. 18). Data normalization was performed using the *edgeR* R package (v. 3.30.2; ref. 18). Specifically, raw counts were normalized to counts per million mapped reads (CPM) and fragments per kilobase million (FPKM), and only genes with a CPM greater than 1 in at least three samples were retained. To identify two groups of tumors with either high or low TLS1 and CD31$^+$plasma cells, we used the classifier described in ref. 19, which is a classification rule based on gene expression signature scores. Briefly, the signature scores have been obtained summarizing the standardized expression levels of CD38 and PECAM1 (CD31) genes into a combined score with zero mean. Tumors were classified as CD38/CD31 signature "Low" if the combined score was smaller than the median signature score and as CD38/CD31 signature "High" vice versa. This classification was applied to expression values of the TCGA HNSC samples with disease-specific survival information ($n = 493$).

Plasma cell fractions have been quantified using CIBERSORT (20) on the TCGA HNSC collection. Briefly, the FPMK expression matrix was uploaded to the CIBERSORT R script (version 1.04) as a mixture file, and CIBERSORT was run in absolute mode with the LM22 signature gene file, 100 permutations, and quantile normalization. In absolute mode, CIBERSORT scales relative cellular fractions into a score that reflects the absolute proportion of each CT in a mixture. Although not expressed as a fraction, the absolute score can be directly compared both between- and within-samples (21). Samples have been divided into two groups based on the median of the absolute scores for the plasma CT.

To evaluate the prognostic value of the CD38/CD31 signature and the plasma cell fraction, we estimated the probabilities of disease-specific survival using the Kaplan–Meier method. The Kaplan–Meier curves were compared using the log-rank (Mantel–Cox) test to confirm these findings. The P value was calculated according to the standard normal asymptotic distribution. Survival analysis was performed using functions of *survival* (version 3.4–0) and *survminer* (version 0.4.9) packages. Multivariate analysis of the association of age, sex pathological stage, histologic grade, plasma cell fraction, and CD38/CD31 signature classification factors with disease-free survival has been performed using the *analyse_multivariate* function of the *survivalAnalysis* R package (version 0.3.0). Forest plot representation of the multivariate analysis results has been obtained using the *forest* function of the *forestplotter* R package (version 0.2.3). To compare the multivariate models, we subtracted the residual deviance of the multivariate model with age, sex pathological stage, histologic grade, and plasma cell fraction (Model 1) from the one with age, sex pathologic stage, histological grade, plasma cell fraction, and CD38/CD31 signature classification (Model 2) and tested this difference against a chi-square distribution with one degree of freedom. All analyses have been done in R 4.2.2.

Data and source availability

The single-cell data table of clustered, annotated CTs with metadata are publicly available in Mendeley (https://data.mendeley.com/datasets/t2yvtwnjx7/1). R scripts created for this study are available at https://github.com/bicciatolab/CODEX_HNSCC.

Additional data analyzed in this study were obtained from the TCGA repository at https://portal.gdc.cancer.gov/. All other raw data generated in this study are available upon request from the corresponding author.

Results

Unsupervised analysis of CODEX data identified 20 CTs in HNSCC tumors

To analyze the immunologic landscape of the HNSCC microenvironment, we performed highly multiplexed microscopy using CODEX on nine surgical specimens of patients with stage II to IV HPV-negative squamous cell carcinoma explants of the oral cavity and oropharynx. We selected and validated a panel of 30 antibodies that allows the discrimination of the main leukocyte subsets infiltrating the tumor and auxiliary markers for tumor, vascular, and stromal cells on OCT embedded flash-frozen sections. To minimize sampling errors, we processed an area of 42.6 mm^2 derived from 121 best-focused stitched 20× fields using the Akoya software and default parameters for quality control, image processing, and cell segmentation. We evaluated the image quality, segmentation, and fluorescence signals on CODEX MAV; we excluded staining artifacts and exported the fluorescence signals, metadata, and spatial coordinates of all cells for processing in R using an *ad hoc* computational framework. For each specimen, we clustered cells at different resolutions (ranging from 0.2. to 0.8) using the centered log-ratio normalized signals and the graph-based clustering algorithm of the Seurat package (22), and we visualized cell clusters on low-dimensional embeddings (t-SNE and UMAP) and on the spatial locations. We associated clusters to CTs at any explored resolution and determined the optimal cluster resolution by visual inspection of marker expression and spatial localization (Supplementary Table S1). We removed staining artifacts ($n = 945$ cells) and merged clusters with similar marker expression profiles and morphological appearance in the tissue. Finally, we generated a merged dataset accounting for 792,349 cells from all nine original specimens. This procedure allowed the definition of 20 phenotypes that, in a UMAP embedding, segregate cells according to the type and not on the specimen of origin (Fig. 1A; Supplementary Table S1). As expected, we identified cytokeratin-positive tumor cells, CD31$^+$epithelial (blood vessel) cells, the main lymphocyte subsets (NK and NKT, CD8$^+$cytotoxic and CD4$^+$ T helper lymphocytes, Treg, B cells, and plasma cells), various populations of myeloid cells (monocytes, M1 and M2 macrophages, neutrophils, Lox$^+$PMN-MDSCs; ref. 23), and hematopoietic stem and progenitor cells (HSPC; ref. 24). In addition, we observed the presence of CD11b$^+$cells expressing the transcription factor FOXP3, IL4Rα, and PDL1 (Fig. 1B; Supplementary Fig. S1A), of a heterogeneous group of myeloid cells expressing PDL1 (PDL1$^+$myeloid), and of CD90$^+$ mesenchymal stem-like stromal cells (25) that surround the neoplastic islands (Fig. 1C; Supplementary Figs. S1B–S1K). Interestingly, the unsupervised clustering identified a population of cells that we named capsule, characterized by low or negative expression of most markers and, when present, contoring the neoplastic nests (Fig. 1C; Supplementary Figs. S1B–S1K). Finally, we compared the frequencies of each cell phenotype in all tumor specimens. Although some CTs were present in all or most patients (e.g., tumor, CD90$^+$ cells, vessel, PMN-MDSC), others (as capsule, NKT, B cells, PDL1$^+$myeloid) were present only in a limited number of samples (Fig. 1D; Supplementary Table S1).

In summary, the analysis of CODEX data revealed a complex tumor microenvironment characterized by multiple cell subsets whose composition differs across tumor specimens.

Flow cytometry and CODEX differentially enumerate the leukocyte subsets in the tumor microenvironment

Flow cytometry is considered the gold standard for enumerating tumor-infiltrating leukocyte subsets. However, this technique requires extensive tissue processing to obtain a single-cell suspension, which may impact the proper representation of different cell subsets. In contrast, CODEX allows the evaluation of the tumor microenvironment using a limited amount of tissue and minimal specimen processing. Despite the limited sample processing, sampling errors may also affect CODEX analyses. We divided each surgical specimen into two parts to compare the performances of flow cytometry and CODEX in enumerating cell subsets. One part was processed and analyzed by multicolor flow cytometry; the other was frozen and evaluated by multiparametric CODEX immunofluorescence microscopy. We identified the major lymphoid and myeloid subsets in FCS Express using similar manual gating strategies on the data originated by the two techniques (Fig. 2A and B). Compared with CODEX, we observed that flow cytometry significantly overestimated the percentage of CD3$^+$ lymphocyte subsets and dramatically underestimated the PMN-MDSC fraction (Fig. 2C and D). We reasoned that these differences might derive from tissue processing that, in the case of tumors preparation for flow cytometry, requires collagenase treatment, filtration, and multiple centrifugations, and we hypothesized that a tissue requiring a simpler processing (i.e., a mouse spleen that requires only mechanical dissociation, filtration, and RBC lysis) could minimize the differences between the two technologies. To test this hypothesis, we enumerated the leukocytes from the spleen and the tumor of mice bearing the 4T1 mammary carcinoma using both flow cytometry and CODEX. Briefly, spleens and tumors were harvested; half of each specimen was snap-frozen and analyzed by CODEX, whereas the other half was processed and analyzed by flow cytometry (Fig. 2E). As observed in human HNSCC, compared with CODEX, flow cytometry overestimated the number of T cells and mMDSC and underestimated the number of PMN-MDSC in the tumor. In contrast, no significant differences were observed when splenic leukocytes were enumerated by the two techniques (Fig. 2E). These results indicate that the low frequency of PMN-MDSC usually found by flow cytometry and scRNA sequencing in mouse and human tumors (26) and the lower PMN-MDSC/m-MDSC ratio observed in tumors (compared with the one observed in circulation or in the spleen) can be attributed to technical artifacts of the tissue preparation.

The HNSCC tumor microenvironment is organized in distinct CNs

Recent data from Nolan's laboratory (10, 11, 14, 15, 27) show that host and neoplastic cells are not randomly distributed in the tumor microenvironment, but they are instead spatially organized in communities called CNs, which may provide the topological and phenotypic architecture for optimal cell–cell interaction and function. An example of a CN is the TLS, a spatial organization of infiltrating immune cells that can positively or negatively modulate antitumor immunity (28). To identify CNs in our samples, we quantified, for each cell of the merged dataset, the phenotype frequencies of its W nearest neighbors and clustered these topological and compositional vectors into K groups with distinctive enrichments of the 20 original CTs (11). To optimize the width of the capturing window W and the number K of CNs, we tested different combinations of neighboring cells (W range 7–200) and CNs (K range 7–14). Finally, we considered a window of $W = 10$ cells and opted for $K = 14$ CNs as this combination (i) maximized the number of CNs with a unique pattern of phenotype frequencies (enrichment score), (ii) minimized the number of CNs enriched in the most abundant CTs (i.e., tumor cells), and (iii) better recapitulated the cellular composition of the extensively characterized TLS neighborhood modeled around plasma cells, CD4$^+$T cells, and stromal cells (29).
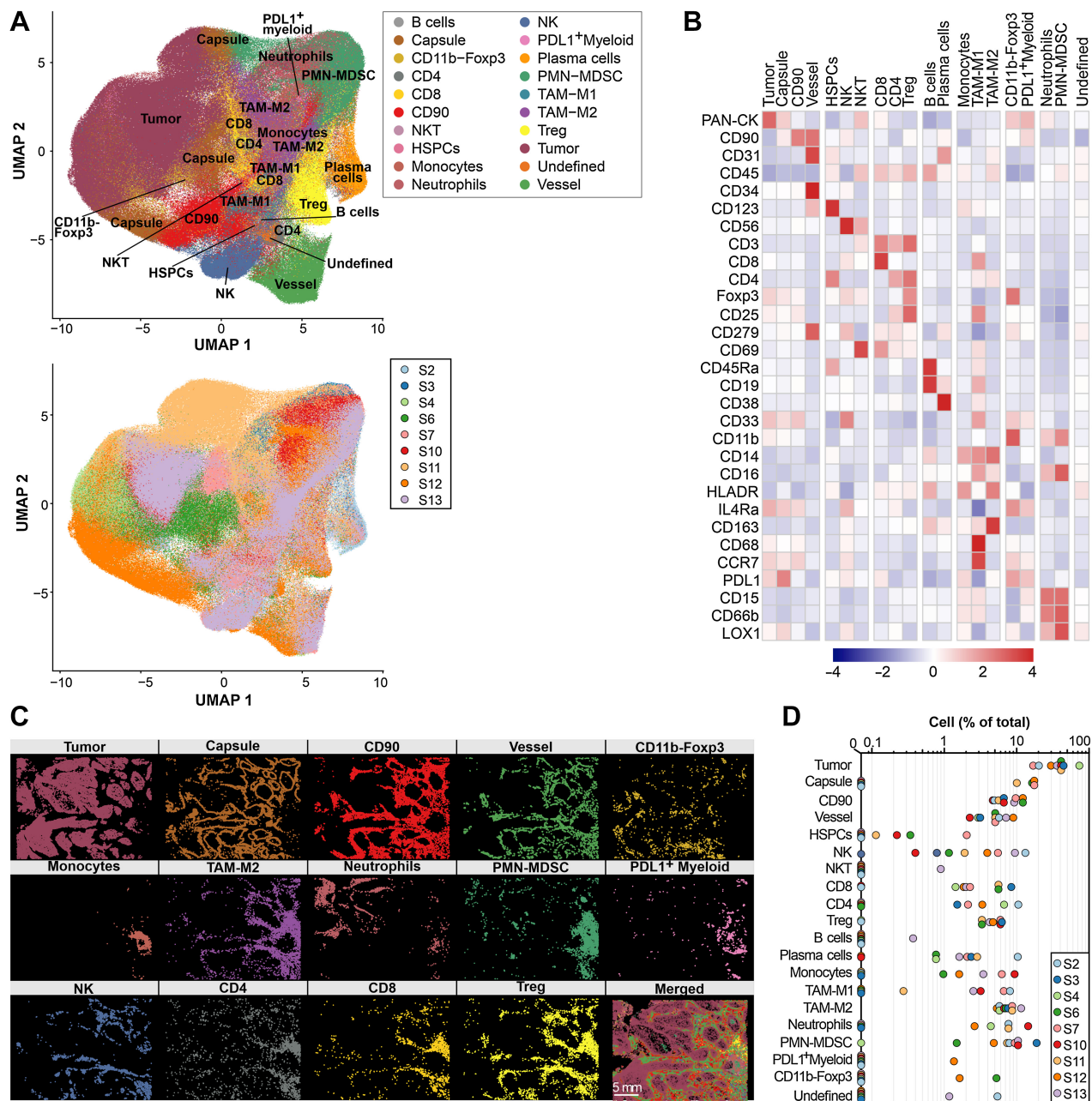
Figure 1.
Unsupervised clustering identified 20 phenotypes in the tumor tissue of patients with HNSCC. CODEX immune fluorescence microscopy data were normalized with the centered log ratio–based normalization and clustered with graph-based clustering to identify CTs. **A,** UMAP of the 792,349 cells from all nine original specimens. Cells are color-coded according to the CT (top) and the specimen of origin (bottom). **B,** Heatmap of mean fluorescence intensities of individual markers across all phenotypes of the 792,349 cells in the merged dataset. **C,** Representative of the spatial distribution of CTs identified in sample s12. The "Merged" panel reports all different CTs together and highlights the presence of spatial tissue structures. **D,** Frequency of CTs across patients. Each dot represents a sample in which the given phenotype has been detected.

After evaluating the CT composition of each CN, we noticed that some clusters were very similar in leukocyte composition (Supplementary Fig. S2; Supplementary Table S1). We then merged the original 14 CNs into 11 CNs with a distinctive phenotype stoichiometry (**Fig. 3A**). We identified a "cold tumor" CN, mainly composed of neoplastic cells; a "hot tumor" CN, characterized by the presence of malignant cells and leukocytes; a "peritumoral" CN made of neoplastic and capsule cells; a "vascular" CN rich in vessel cells; and a "stromal CN" composed of CD90+MSC-like cells (**Fig. 3A**). In addition, the analysis revealed the presence of NK-rich, M1-rich, neutrophils-rich, and MDSC-rich CNs characterized by the high frequency of NK, TAM-M1, LOX1−neutrophil, and LOX1+PMN-PMN-
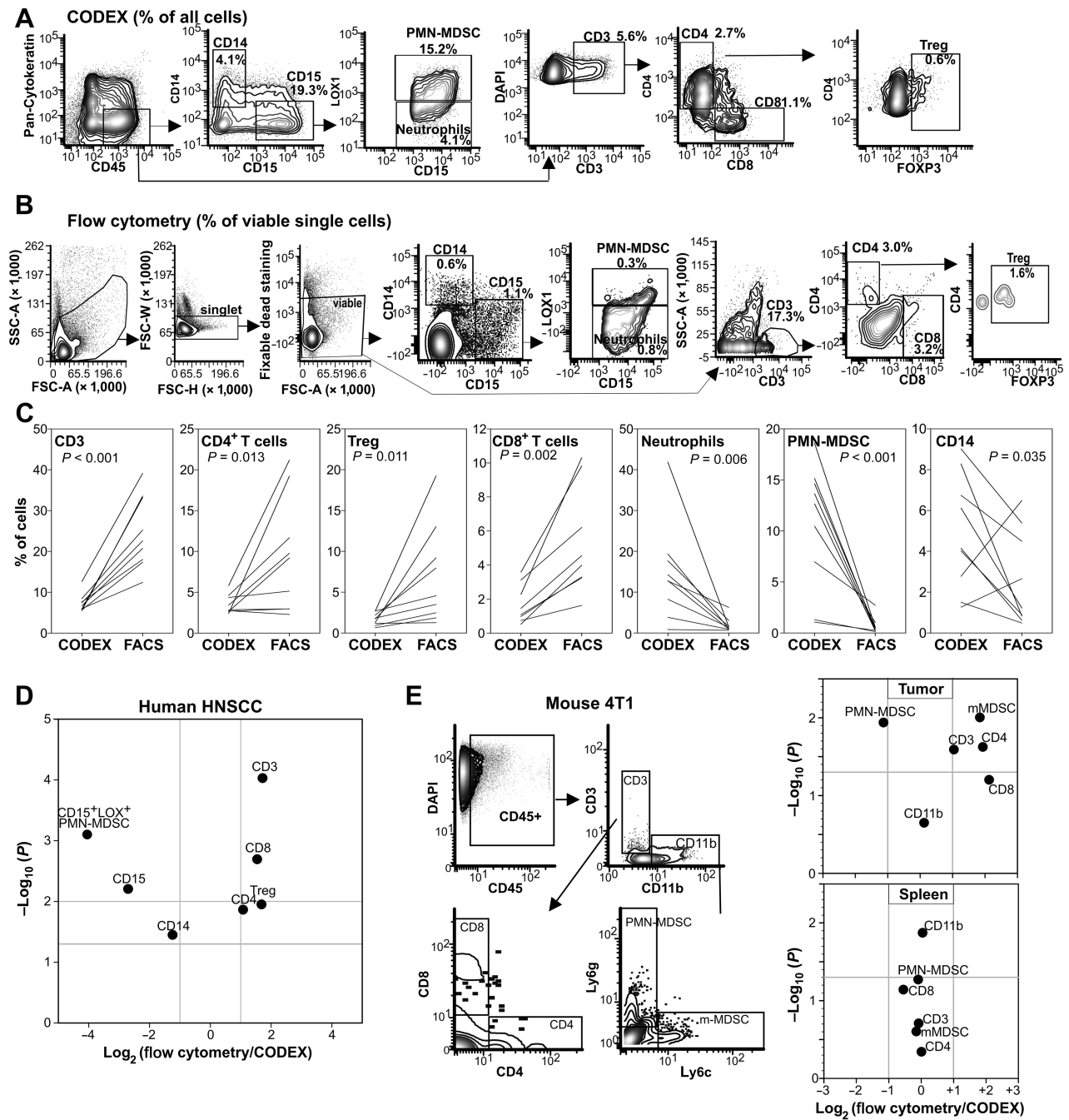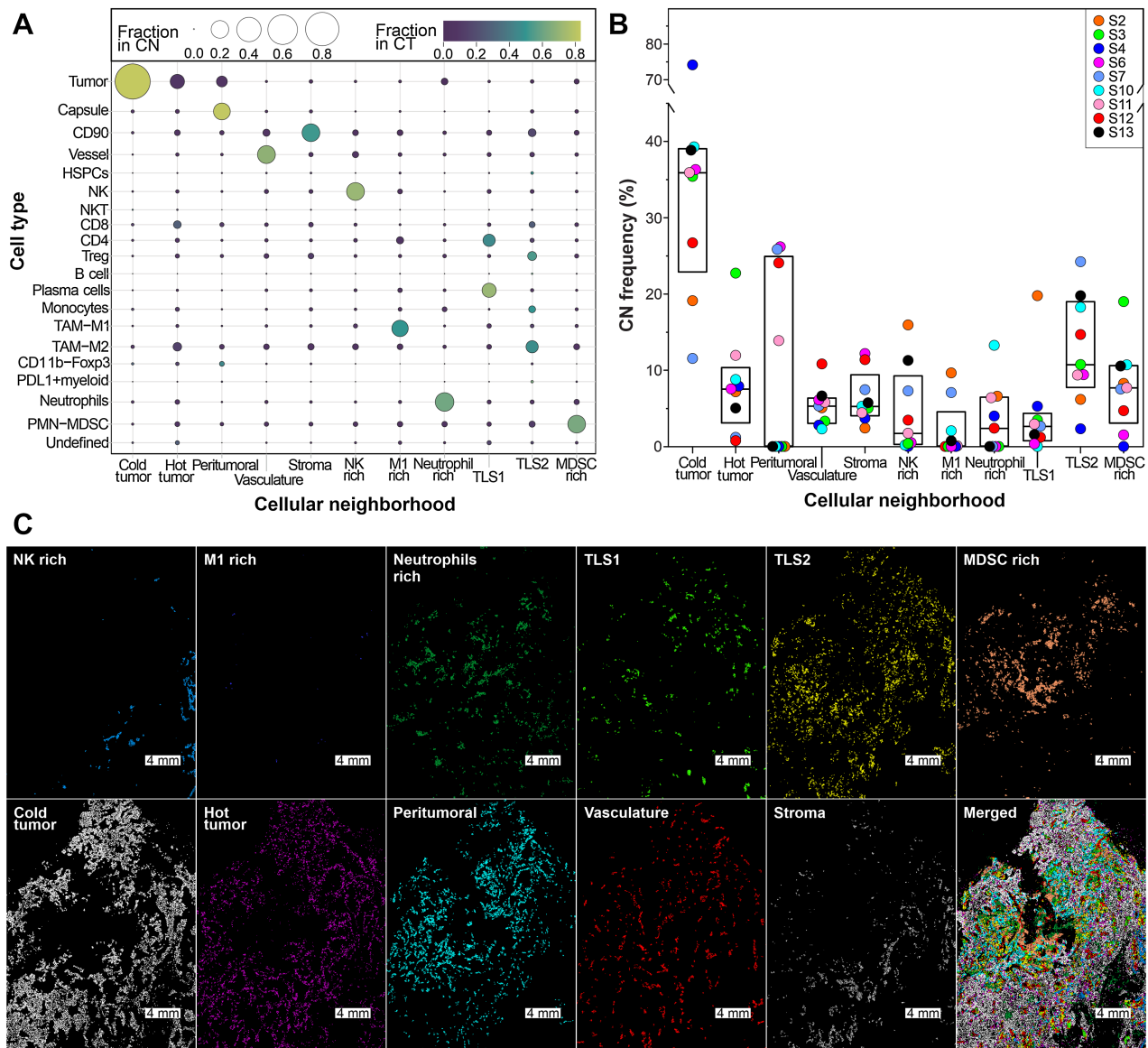
Figure 2.
CODEX and flow cytometry differently estimate the frequency of TILs. Surgical specimens were processed within 2 hours of resection. Part of each specimen was snapped frozen for subsequent CODEX analysis, and part was processed for immediate multicolor flow cytometry. **A** and **B,** Signals were acquired at a single-cell level, and a similar manual gating strategy was employed to identify the frequency of the indicated leukocyte subsets acquired by flow cytometry (**A**) and CODEX (**B**). The percentage shows the frequencies within viable single cells in flow cytometry analysis or the frequencies of all cells in CODEX analysis. **C,** Phenotype frequencies estimated by the two techniques (P values from paired t test). **D,** Volcano plot comparing the frequencies of the indicated leukocyte subsets measured by multicolor flow cytometry over CODEX on the same specimens. **E,** Splenic and tumor-associated leukocytes from mice (n = 3) bearing the 4T1 mammary carcinoma were enumerated by flow cytometry and CODEX using the shown gating strategy.

MDSCs, respectively. Finally, different leukocyte subsets characterized two distinct CNs, that is, TLS1, rich in plasma cells and CD4⁺Foxp3⁻ lymphocytes, and TLS2, characterized by the presence of M2-like macrophages and Tregs. We assessed the CN frequencies in the nine

tissue samples and noticed that cold tumor and TLS2 are the most abundant CNs, followed by peritumoral (when present), hot tumor, MDSC-rich, and stroma CNs (**Fig. 3B**; Supplementary Table S1). We observed substantial variation in the frequency of the other CNs: five

Figure 3.
Discrete CNs can be identified in the HNSCC microenvironment. **A,** CT composition of the 11 merged CNs. The size of the circles indicates the frequency of a given CT in a CN over the total cell number in the same CN. The color of the circles indicates the fraction of a given CT in a particular CN over the total number of that CT in all samples. **B,** Frequencies of the CNs in the nine tissue samples (each color represents a patient). **C,** Representative of the spatial distribution of CN in an example specimen (i.e., sample s11).

samples lacked the peritumoral CN, and four were deficient in the neutrophil-rich and the M1-rich CNs (**Fig. 3B**; Supplementary Table S1). We then evaluated whether some correlation existed among the CN frequencies. We found that the stroma CNs negatively and positively correlated with the TLS1 ($r = -0.80$; $P = 0.006$) and the peritumoral CNs ($r = 0.73$; $P = 0.020$), respectively. Similarly, we observed a negative correlation ($r = -0.67$; $P = 0.043$) between the frequency of NK and cold tumor CNs (Supplementary Fig. S2B). As expected, the hot tumor and the peritumoral CNs separated the cold tumor CN from all other CNs (**Fig. 3C**; Supplementary Figs. S2C–S2T). Generally, the MDSC-rich CN and the stroma occupied the space between the neoplastic nests and contained the other CNs. In

most samples, the NK-rich CN was located distantly from the tumoral CNs, whereas neutrophil-rich CN, when present, was in proximity and interacted with the neoplastic cells (**Fig. 3C**; Supplementary Figs. S2C–S2T).

Finally, we reasoned that the CNs could differ not only in the CT composition but also in the interaction between CTs. Thus, we quantified the pairwise cell–cell contacts for CTs within CNs (**Fig. 4**; Supplementary Fig. S3). We observed a higher degree of homotypic cell–cell contacts in some CNs (e.g., neoplastic cells with neoplastic cells, CD90+cells with CD90+cells, etc.) whereas, in other neighborhoods, cell–cell contacts were more heterotypic. For instance, in the peritumoral CN, capsule cells interacted with other capsule cells but
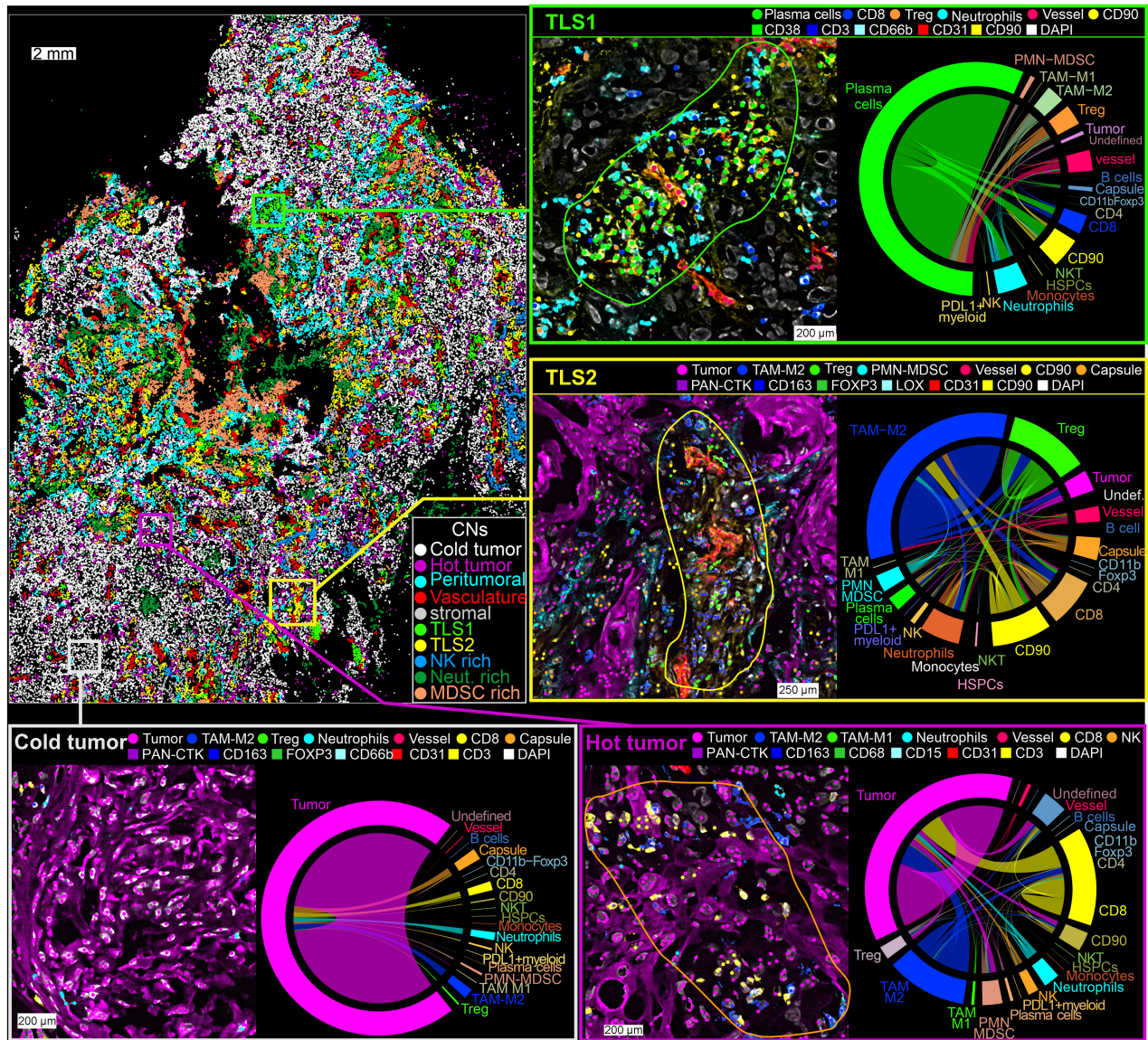
Figure 4.
CNs can be differentiated according to cell-to-cell contacts. Representative of cell–cell interaction in CNs of sample s11. CT and CNs were imported in CODEX MAV, and their locations overlaid to the multicolor picture. A ×20 magnification is shown for TLS1, TLS2, hot tumor, and cold tumor CNs. Key makers associated with a particular CT are shown, and different CTs are indicated by filled circles. The circle plots on the right indicate the contacts between CTs in the given CN. Circle plots of cell–cell contacts within each CN for all cells of the merged dataset are reported in Supplementary Fig. S3.

also with tumor and leukocyte cells, forming a barrier between these latter CTs (Supplementary Fig. S3). We found the largest number of heterotypic interactions in the TLS1, TLS2, and hot tumor CNs (**Fig. 4**; Supplementary Fig. S3). Plasma cells dominated the cellular interactions in the TLS1 neighborhood. These cells interacted with other plasma cells but also with CD90$^+$MSC-like cells, vessel cells, and both CD4$^+$ and CD8$^+$ T cells (**Fig. 4**). This finding is consistent with the type of TLS described in many cancer types and is generally associated with a good prognosis (30). M2 macrophages dominated the interactions within the TLS2. These macrophages interacted with Treg, PMN-MDSC, CD4, and CD8 T cells, suggesting a phenotype that may inhibit antitumor immunity (**Fig. 4**).

Because the same CTs can be found (although at different frequencies) in both TLS1 and TLS2, we evaluated whether cell–cell interactions for a given CT were influenced by the cell neighborhood it belonged to (i.e., did a plasma cell interact with different CTs if it belongs to TLS1 or TLS2?). Thus, we compared the relative cell–cell contact frequencies (11) for all phenotypes in different pairs of CNs (e.g., TLS1 and TLS2; hot and cold tumor) across all patients.

We found that, in TLS1, plasma cells interacted mostly with other plasma cells and were contacted by NK, CD90$^+$stromal, and PNM-MDSC cells (**Fig. 5A**; Supplementary Table S1). Instead, in TLS2, M2 macrophages, PMN-MDSC, CD8 T lymphocytes, Treg, vessel cells, and CD90 cells interacted significantly more with M2 macrophages. In
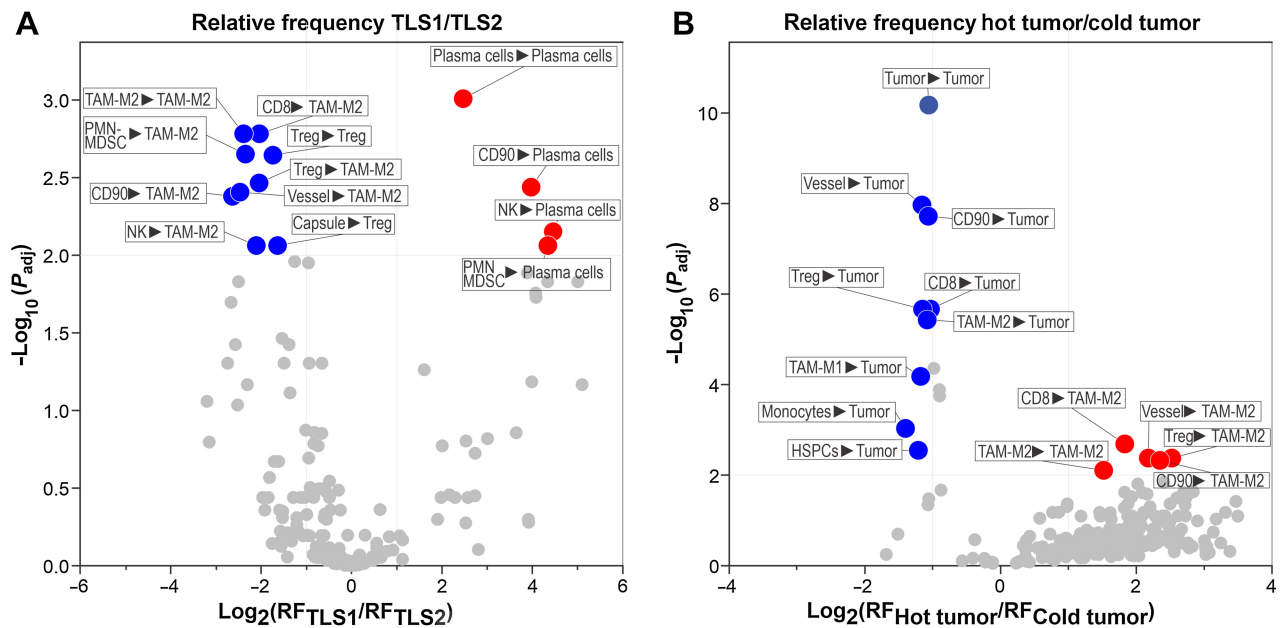
Figure 5.
Significant differences in cell–cell interactions across CNs. **A,** Comparison of the relative cell–cell contact frequency in TLS1 over TLS2. Cell–cell interactions with RF absolute fold change ≥2 and adjusted P value ≤0.01 have been considered significant and are highlighted in blue (smaller relative contact frequency in TLS1) and red (larger relative contact frequency in TLS1). **B,** Same as in **A** for comparing hot tumor over cold tumor CNs.

TLS2, additional significant contacts were those between Tregs with other Tregs and capsule cells with Tregs (**Fig. 5A**; Supplementary Table S1).

In the hot tumor CN, approximately one-third of the neoplastic cell contacts were with leukocyte subsets such as neutrophils, CD4 and CD8 T cells, but also with M2 macrophages and PMN-MDSC (**Fig. 4**; Supplementary Fig. S3). In contrast, in the cold tumor CN, most neoplastic cell interactions were homotypic (**Fig. 4**; Supplementary Fig. S3). The differential analysis of relative contact frequencies confirmed these observations and indicated that, in the cold tumor CN, neoplastic cells form a significantly higher number of homotypic interactions and that the few leukocytes interact primarily with tumor cells (**Fig. 5B**; Supplementary Table S1). Conversely, in the hot tumor neighborhood, CD8$^+$ Tcells, vessels, Treg, M2 macrophages, and stromal cells interact significantly more with M2 macrophages (**Fig. 5B**; Supplementary Table S1).

These analyses revealed a complex tumor microenvironment in which CTs are organized in spatially distinct units that are different in cell composition, cell–cell interactions, and colocalization. These CNs may be functionally important for tumor growth and immune escape or the generation of effective antitumor immunity.

The architecture of the tumor microenvironment correlates with clinical parameters and prognosis

The composition of the tumor microenvironment is a major determinant of patients' prognosis and tumor clinical and pathological features. Hence, we would expect that patterns of CTs and CNs in spatial regions are differently coupled and correlated, in patients, with different clinical phenotypes (e.g., recurring or not after surgery, presenting or not perineuronal invasion, or with and without primary tumor at surgery). As demonstrated by Schurch and colleagues, tensor decomposition can reveal how cell phenotypes and CNs are combined

into distinct tissue modules in groups of patients with different clinical characteristics (11). We, therefore, constructed a joint composition matrix quantifying the frequency of each CT in each CN and collected these matrices in tridimensional tensors. Of the 9 analyzed patients, 4 recurred within a year from surgery (median recurrence time 0.65 years, range 0.35–0.92 years), 4 did not recur during the observation time (median observation time 2 years, range 1.2–2.1 years), and 1 patient was uncontactable after surgery. To determine whether any characteristic of the tumor microenvironment was associated with early (<1 year) tumor recurrence after salvage surgery, we constructed a 3D tensor for those patients that recurred after surgery and one for those that did not recur (**Fig. 6A**). Then, we applied non-negative Tucker tensor decomposition to the tensor of patients' joint CT–CN matrices in recurrent and non-recurrent patient groups separately (11, 31). Tensor decomposition analysis returned a set of eight CN and eight CT modules (dashed inner rectangles in **Fig. 6B**) that shaped three different tissue modules in patients recurring and not recurring after surgery (solid outer rectangles in **Fig. 6B**). Graphically, we represented the contribution of each CT and CN to each tissue module based on its value in the corresponding slice of the core tensor and used this value for color shading. Elements with a value lower than 0.1 in the corresponding slice of the core tensor were not deemed to contribute to a given tissue module and were greyed out (**Fig. 6B**). In both recurrent and non-recurrent patients, the first tissue module was characterized by neoplastic, protumoral leukocytes, and stroma CN and CT components. Similarly, factors belonging to the stroma and stroma–tumor interface contributed to defining the CN and CT components in the second modules for both groups of patients, whereas the vasculature CN weighed in the definition of the second tissue module in those patients that recurred but did not contribute to the same module in non-recurrent patients. Strikingly, the third tissue module was defined by different CN and CT components in recurrent and non-recurrent
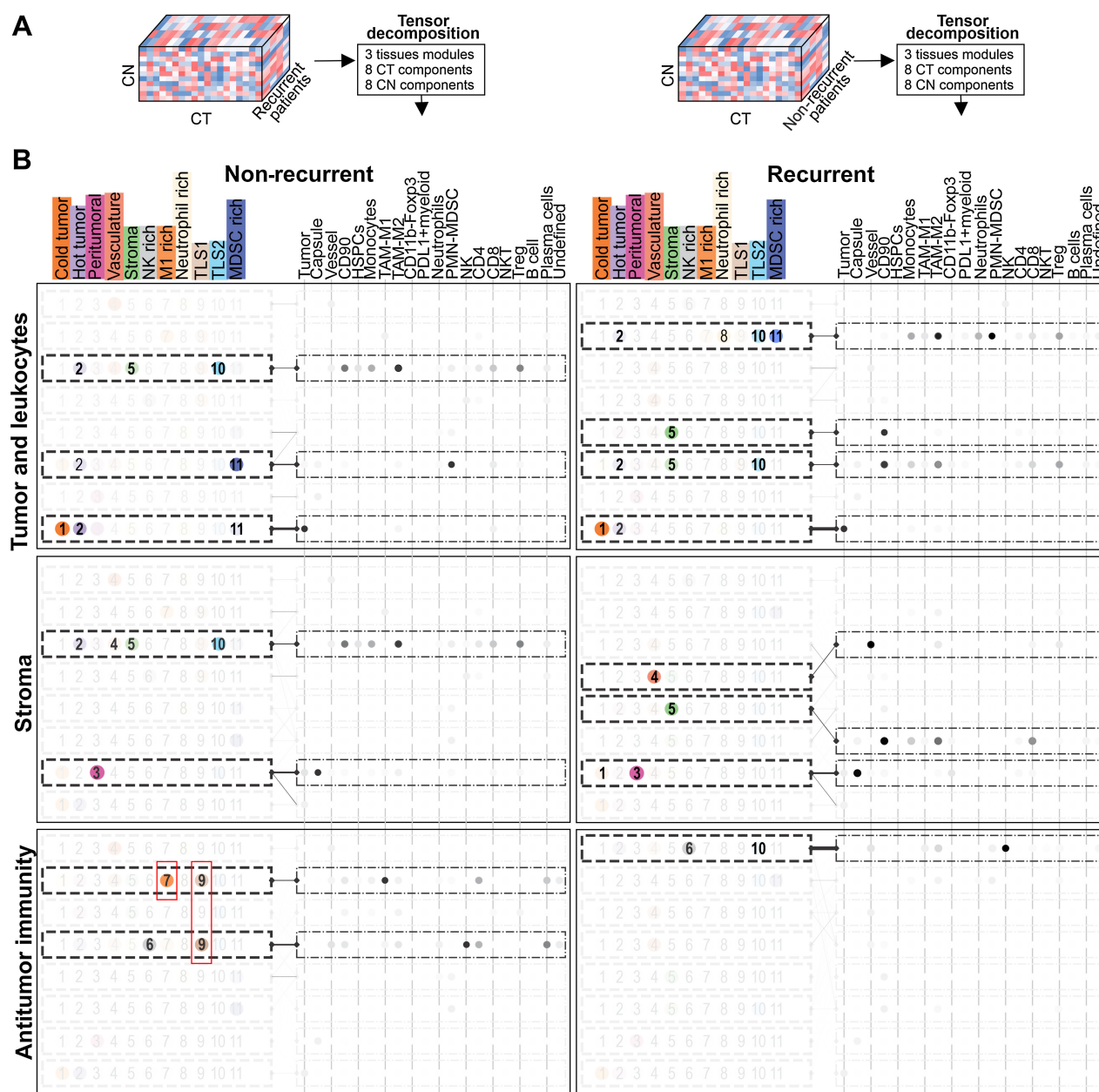
Figure 6.
Tensor decomposition analysis discriminates the tumor microenvironment of recurrent and nonrecurrent patients. **A,** Schematic of the tensor decomposition analysis. **B,** Decomposition results for recurrent (recurrence within 365 days from surgery) and nonrecurrent patients (no recurrence in the first year). Tissue modules containing interacting CN and CT components were named according to the CTs and CNs that weighed the most in each module. Components with small weights are grayed out. The strength of the interaction between CN and CT components is proportional to the width of the connecting lines. The intensity of the color is proportional to the weight with which a particular CN or CT contributes to the component.

patients: while in recurrent patients, NK CT weighed the most in the CT and CN components, in non-recurrent patients, the NK, M1 macrophages, and TLS1 weighed in the definition of the CN components whereas M1 macrophages, NK, CD4 T cells and plasma cells defined the CT component (**Fig. 6B**). In accordance with this finding, the TLS1 CN was more frequent in patients whose tumor did not recur (**Fig. 7A**) and its presence significantly correlated with the absence of recurrence (Supplementary Fig. S4A) and was predictive of recur-

rence-free survival (**Fig. 7B**). Similarly, plasma cells weighed in the tissue module of no recurrent tumors (**Fig. 6B**), were associated with the absence of recurrence (Supplementary Fig. S4B) and predicted recurrence-free survival (**Fig. 7C**). Of note, besides perineuronal invasion ($P = 0.0042$), no other clinical or pathological parameters (e.g., smoking history, age, sex, history of previous HNSCC) predicted recurrence after surgery in univariate analysis. Because of the limited amounts of patients, multivariate analysis could not be performed.
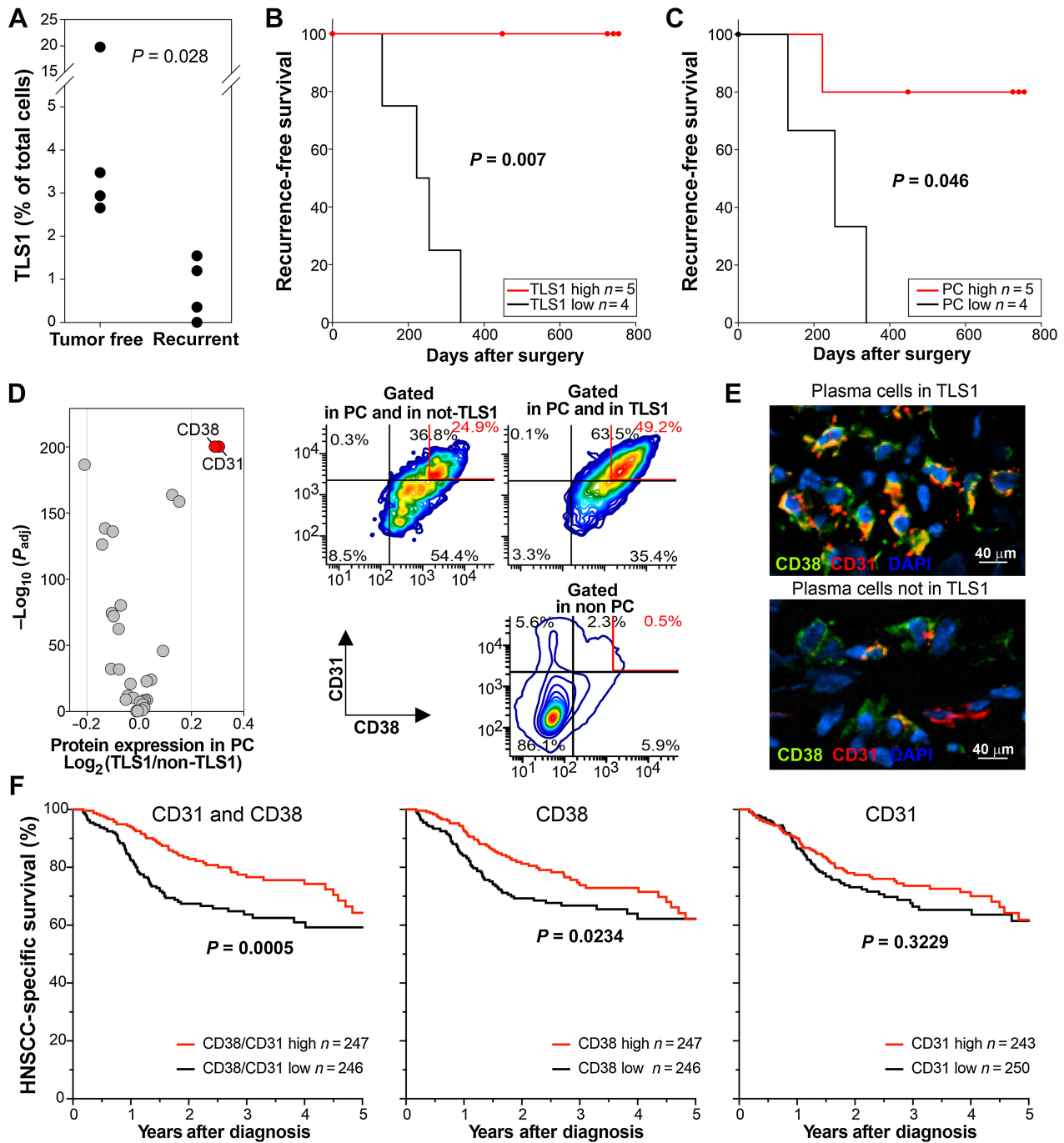
Figure 7.
TLS1 and CD31[high]CD38[high] plasma cells' frequency correlates with a better prognosis in patients with HNSCC. **A,** TLS1 frequency in recurrent and nonrecurrent patients (P value from Mann–Whitney rank sum test). **B,** Survival analysis according to TLS1 frequency. Patients were stratified according to the median frequency of cells belonging to the TLS1 CN. **C,** Survival analysis according to plasma cell (PC) frequency. Median was used to dichotomize the patient's group. **D,** Differential marker expression analysis of plasma cells in TLS1 compared with plasma cells in all other CNs (left). The contour plots on the right were generated from the merged raw data database containing the CN and CT identifiers. The expression of CD31 and CD38 was plotted after gating on (i) PC belonging to the TLS1 CN, (ii) plasma cells not belonging to the TLS1 CN, or (iii) nonplasma cells. **E,** Example of plasma cells in TLS1 or in other CNs. **F,** Survival analysis of TCGA HNSC dataset. Samples were stratified in high and low groups according to the CD38 and/or CD31 signature score obtained by summarizing the standardized expression levels of CD38 and PECAM1 (CD31) genes into a combined score with zero mean. Log-rank P values are shown.

Interestingly, we also found that the frequency of cells from the cold tumor CN positively correlated with the perineural invasion (PNI) and the depth of invasion (DOI). In contrast, the peritumoral and MDSCs-rich CNs correlated with the tumor size and previously irradiated fields, respectively. We also observed a significant negative correlation between patient smoking history and the frequency of cells from the NK-rich CN (Supplementary Fig. S4A).

Given the small number of patients evaluated in this study, we investigated whether we could identify distinctive markers of plasma cells in TLS1 and used them to interrogate larger collections of patients with HNSCC. Hence, we performed a differential protein expression analysis between plasma cells in the TLS1 and those belonging to other CNs. We found that plasma cells in TLS1 expressed significantly higher levels of the CD38 and CD31 proteins (**Fig. 7D** and **E**). These markers define a subset of long-lived IgG-secreting plasma cells (32) usually found in primary and secondary lymphoid organs but not in circulation (33–35).

To evaluate whether the expressions of CD38 and CD31 in the tumor microenvironment had a prognostic value, we classified HNSCC tumors from the TCGA dataset based on the single and combined expression of these two genes. We found that higher transcriptional levels of CD38 and CD31 significantly correlated with better disease-specific survival (**Fig. 7F**), whereas the expression of either CD38 or CD31 alone was not or only slightly predictive ($P > 0.01$). Of note, the expression level of the CD38/CD31 signature was the only feature significantly ($P = 0.0006$) associated with a better prognosis in multivariate analysis with age, sex pathologic stage, and histological grade (Supplementary Fig. S4C). A similar significant ($P = 0.0055$) association was found when using as covariate the plasma cell fraction inferred using CIBERSORT (20) deconvolution algorithm (Supplementary Figs. S4D and S4E). Adding CD38/CD31 signature to this multivariate analysis conferred additional predictive value (Supplementary Fig. S4E). Similar correlations between the CD38/CD31 signature and a better prognosis were found using the Kaplan–Meier Plotter resource (www.kmplot.com; ref. 36) in renal clear cell carcinoma (HR, 0.42; $P = 5 \times 10^{-9}$), lung adenocarcinoma (HR, 0.65; $P = 0.0061$), and thymoma (HR, 0; $P = 0.0001$).

Taken together, these data support the notion that the spatial organization of the tumor microenvironment is essential for cell function and prognosis and points to TLS1 and CD31^{high}CD38^{high} plasma cells as important predictors of tumor recurrence.

Discussion

Previous studies analyzed the interaction between the tumor and the immune system well and showed the pivotal role of different leukocyte subsets in neoplastic cell growth, recurrence, and metastases (37). However, only a few retrospective studies evaluated the tumor microenvironment globally, considering cell-to-cell interactions, spatial context, and relationships between CT phenotype and surrounding cells (11, 12). These studies revealed that the tumor is a highly organized tissue in which leukocytes interact within structures or CNs (11, 12). The presence, frequency, and composition (in terms of cell phenotypes) of these micro-niches might be relevant in shaping the optimal environment for leukocyte interaction and function and, ultimately, in determining tumor characteristics.

We performed the first prospective study in patients with HNSCCs to evaluate whether features of the fine architecture of the tumor microenvironment are associated with the risk of recurrence after salvage surgery. We analyzed the HNSCC samples using the CODEX-based multiparametric immunofluorescence method with an opti-

mized 30 markers antibody panel and an *ad hoc* computational framework. Using a manual gating strategy, we compared the performances of CODEX in quantifying the frequency of the main leukocyte subsets with those of multicolor flow cytometry. We found that, compared with CODEX, flow cytometry overestimated the T-cell frequency and dramatically underestimated the polymorphonucleate populations and, in particular, the PMN-MDSCs. MDSCs emerged as a crucial population that inhibits tumor immunity and promotes tumor growth and metastases (38). Interestingly, different studies showed that the PMN-MDSC/M-MDSC ratio differs in tumors and in the periphery (26). These observations could indicate a real biological property (i.e., a preferential migration of M-MDSC to the tumor or their enhanced survival) or could derive from a selective loss of PMN-MDSC during isolation. Because PMN-MDSC and M-MDSC are functionally distinct (38), this conundrum is clinically essential for developing effective immunotherapies. Our data in mice and humans support the notion that the conventional techniques significantly underestimate PMN-MDSC, suggesting caution in determining the frequency of leukocyte subsets in the tumor microenvironment following enzymatic and mechanical processing of the tissue (**Fig. 2**).

Because cell phenotyping by hand-gating is subjective, sensitive to segmentation noise, and cumbersome with high-dimensional data (39, 40), we analyzed the normalized CODEX data with a clustering-based approach followed by manual annotation and identified 20 distinct CTs (**Fig. 1**). Besides the expected lymphoid and myeloid subsets present at various ratios across the analyzed tumors, this analysis revealed the presence of CD90$^+$ stromal cells, a population of cells lining around the neoplastic nests that we called capsule cells, and a rare population of CD11b$^+$Foxp3$^+$ cells, positive for PDL1 and IL4Rα, whose existence and biological relevance is however still controversial (41–43). The abundance of CD90$^+$ cells between the neoplastic island confirmed previous studies indicating CD90$^+$ mesenchymal stem cells as a major contributor to HNSCC stroma (25). CD90$^+$ mesenchymal stem cells drove glioma, colon carcinoma, breast cancer, and HNSCC progression by increasing neoplastic cell proliferation, migration, and adhesion (44).

The identification of the capsule cells was an unexpected finding of this study. These cells were negative for all tested lineage markers but delineated the neoplastic islets in 4 of the 9 patients and were positive for PDL1, suggesting a possible protective role from the immune response. Similar structures composed of PDL1$^+$ cancer associated fibroblasts (CAF) have been reported in breast cancer and shown to correlate with a better prognosis (45). Other studies, however, indicate CAFs as promoters of tumor-induced tolerance, neo-angiogenesis, and cancer stemness (46). Further studies are needed for a finer characterization of the capsule cells, their attribution to the CAF family, and the elucidation of their role in HNSCC progression.

As mentioned above, one of the most important findings of this paper is the identification of CNs as spatial entities accounting for the complexity of the HNSCC tumor microenvironment. Although some CNs were characterized by the homotypic interaction of a particular CT (e.g., MDSC-rich, M1-rich, NK-rich CNs), others were defined by the presence and interaction of different CTs. We focused on two CNs that we named TLS1 and TLS2. The TLS2 CN was characterized by M2 macrophages interacting with Treg, PMN-MDSC, stromal cells, vasculature, and different effector leukocytes such as CD4 and CD8 T cells, NK, and neutrophils. To our knowledge, a similar structure has been described only by a few reports and was shown to correlate with a worse prognosis in patients with breast and pancreatic cancers (47, 48).

In contrast, the TLS1 neighborhood was dominated by plasma cells that interact with CD4$^+$ and CD8$^+$ effector T cells, monocytes, neutrophils, and M1 macrophages. This structure is consistent with a tertiary lymphoid organ mimicking the lymph nodal germinal centers and correlating with a good prognosis in many human malignancies (49). In agreement with previous studies in melanoma, colon carcinoma, and breast cancer (49), our data indicate that frequency TLS1 and plasma cells are good prognostic factors in HNSCC (**Figs. 6** and **7**). We hypothesize that phenotype and possibly the function of a particular CT depend on the spatial cellular context in which it resides. Indeed, differential protein expression analysis on plasma cells from TLS1 compared with plasma cells from other neighborhoods revealed the high expression of the CD38 and CD31 markers (**Fig. 7**). Interestingly, the co-expression of both CD38 and CD31 discriminates patients with a better prognosis in the TCGA HNSC dataset and adds prognostic value to deconvolution models using the plasma cells signature to classify the patients. This finding supports the notion that cell phenotype and function depend on the spatial context and highlights the power of spatial proteomic analysis. Of note, the expression of CD31 and CD38 is associated with a population of long-lived plasma cells usually found in the bone marrow and the lymph nodes (50). Long-lived plasma cells provide a lifelong production of antibodies in mice and humans (51) and are suspected of driving persistent inflammation in autoimmune diseases (51). Although these cells have not yet been described in the tumor microenvironment, their presence might indicate the secretion of high-affinity antibodies that facilitate antitumor immunity by antibody-mediated cell cytotoxicity or by favoring antigen uptake and cross-presentation by antigen-presenting cells (51).

In summary, our study highlights the power of CODEX immune fluorescence in immuno-monitoring, reveals the extraordinary complexity of the tumor microenvironment of HNSCC patients, and supports the notion that TLS1 and long-lived plasma cells are essential determinants of antitumor immunity and tumor progression. However, it is important to note that despite the large number of images ($\sim$100,000) we evaluated, this explorative prospective study is limited by the small number of patients. Thus, the key finding needs to be further assessed in a larger group of patients, possibly using a smaller number of markers or technologies that reduce the data size to increase the speed of computational analysis.

Authors' Disclosures

No disclosures were reported.

Authors' Contributions

D.T. Weed: Conceptualization, resources, writing–review and editing. **S. Zilio:** Investigation. **C. McGee:** Investigation. **B. Marnissi:** Resources, investigation. **Z. Sargi:** Resources. **E. Franzmann:** Resources. **G. Thomas:** Resources. **J. Leibowitz:** Resources. **E. Nicolli:** Resources. **D. Arnold:** Resources. **S. Bicciato:** Conceptualization, resources, data curation, software, formal analysis, supervision, funding acquisition, methodology, writing–original draft, project administration, writing–review and editing. **P. Serafini:** Conceptualization, resources, data curation, formal analysis, supervision, funding acquisition, investigation, methodology, writing–original draft, project administration, writing–review and editing.

Acknowledgments

Note

Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

References

1. Yao Y, Yan Z, Lian S, Wei L, Zhou C, Feng D, et al. Prognostic value of novel immune-related genomic biomarkers identified in head and neck squamous cell carcinoma. J Immunother Cancer 2020;8:e000444.
2. Weed DT, Walker G, De La Fuente AC, Nazarian R, Vella JL, Gomez-Fernandez CR, et al. FOXP3 subcellular localization predicts recurrence in oral squamous cell carcinoma. PLoS One 2013;8:e71908.
3. So YK, Byeon SJ, Ku BM, Ko YH, Ahn MJ, Son YI, et al. An increase of CD8(+) T cell infiltration following recurrence is a good prognosticator in HNSCC. Sci Rep 2020;10:20059.
4. Weed DT, Vella JL, Reis IM, De la Fuente AC, Gomez C, Sargi Z, et al. Tadalafil reduces myeloid-derived suppressor cells and regulatory T cells and promotes tumor immunity in patients with head and neck squamous cell carcinoma. Clin Cancer Res 2015;21:39–48.
5. Li F, Zhao Y, Wei L, Li S, Liu J. Tumor-infiltrating Treg, MDSC, and IDO expression associated with outcomes of neoadjuvant chemotherapy of breast cancer. Cancer Biol Ther 2018;19:695–705.
6. Fattori S, Roux H, Connen E, Robert L, Gorvel L., Le Roy A, et al. Therapeutic targeting of tumor-infiltrating regulatory T cells in breast cancer. Cancer Res 2022;82:3868–79.
7. Altorki NK, Markowitz GJ, Gao D, Port JL, Saxena A, Stiles B, et al. The lung microenvironment: an important regulator of tumour growth and metastasis. Nat Rev Cancer 2019;19:9–31.
8. Finotello F, Eduati F. Multi-omics profiling of the tumor microenvironment: paving the way to precision immuno-oncology. Front Oncol 2018;8:430.
9. Denisenko E, Guo BB, Jones M, Hou R, de Kock L, Lassmann T, et al. Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. Genome Biol 2020; 21:130.
10. Goltsev Y, Samusik N, Kennedy-Darling J, Bhate S, Hale M, Vazquez G, et al. Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. Cell 2018;174:968–81.
11. Schürch CM, Bhate SS, Barlow GL, Phillips DJ, Noti L, Zlobec I, et al. Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. Cell 2020;182:1341–59.
12. Zhang W, Li I, Reticker-Flynn NE, Good Z, Chang S, Samusik N, et al. Identification of cell types in multiplexed in situ images by combining protein expression and spatial information using CELESTA. Nat Methods 2022;19:759–69.
13. Black S, Phillips D, Hickey JW, Kennedy-Darling J, Venkataraaman VG, Samusik N, et al. CODEX multiplexed tissue imaging with DNA-conjugated antibodies. Nat Protoc 2021;16:3802–35.
14. Phillips D, Matusiak M, Gutierrez BR, Bhate SS, Barlow GL, Jiang S, et al. Immune cell topography predicts response to PD-1 blockade in cutaneous T cell lymphoma. Nat Commun 2021;12:6726.
15. Shekarian T, Zinner CP, Bartoszek EM, Duchemin W, Wachnowicz AT, Hogan S, et al. Immunotherapy of glioblastoma explants induces interferon-γ; responses and spatial immune cell rearrangements in tumor center, but not periphery. Sci Adv 2022;8:eabn9440.

16. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995; 57:289–300.

17. Kossaifi J, Panagakis Y, Anandkumar A, Pantic M. TensorLy: tensor learning in python. J Mach Learn Res 2019;20:1–6.

18. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res 2016;44:e71.

19. Adorno M, Cordenonsi M, Montagner M, Dupont S, Wong C, Hann B, et al. A Mutant-p53/Smad complex opposes p63 to empower TGFbeta-induced metastasis. Cell 2009;137:87–98.

20. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015; 12:453–7.

21. Sturm G, Finotello F, Petitprez F, Zhang JD, Baumbach J, Fridman WH, et al. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. Bioinformatics 2019;35:i436–i45.

22. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell 2021;184:3573–87.

23. Condamine T, Dominguez GA, Youn JI, Kossenkov AV, Mony S, Alicea-Torres K, et al. Lectin-type oxidized LDL receptor-1 distinguishes population of human polymorphonuclear myeloid-derived suppressor cells in cancer patients. Sci Immunol 2016;1:aaf8943.

24. Zilio S, Bicciato S, Weed D, Serafini P. CCR1 and CCR5 mediate cancer-induced myelopoiesis and differentiation of myeloid cells in the tumor. J Immunother Cancer 2022;10:e003131.

25. Liotta F, Querci V, Mannelli G, Santarlasci V, Maggi L, Capone M, et al. Mesenchymal stem cells are enriched in head neck squamous cell carcinoma, correlates with tumour size and inhibit T-cell proliferation. Br J Cancer 2015;112: 745–54.

26. Kumar V, Patel S, Tcyganov E, Gabrilovich DI. The nature of myeloid-derived suppressor cells in the tumor microenvironment. Trends Immunol 2016;37:208–20.

27. Phillips D, Schürch CM, Khodadoust MS, Kim YH, Nolan GP, Jiang S. Highly multiplexed phenotyping of immunoregulatory proteins in the tumor microenvironment by CODEX tissue imaging. Front Immunol 2021;12:687673.

28. Rodriguez AB, Engelhard VH. Insights into tumor-associated tertiary lymphoid structures: novel targets for antitumor immunity and cancer immunotherapy. Cancer Immunol Res 2020;8:1338–45.

29. Sautès-Fridman C, Petitprez F, Calderaro J, Fridman WH. Tertiary lymphoid structures in the era of cancer immunotherapy. Nat Rev Cancer 2019;19:307–25.

30. Kang W, Feng Z, Luo J, He Z, Liu J, Wu J, et al. Tertiary lymphoid structures in cancer: the double-edged sword role in antitumor immunity and potential therapeutic induction strategies. Front Immunol 2021;12:689270.

31. Kim YD, Choi S. Nonnegative tucker decomposition. 2007;2007:1–8.

32. Itoua Maïga R, Bonnaure G, Tremblay Rochette J, Néron S. Human CD38^{hi}CD138^{+} plasma cells can be generated in vitro from CD40-activated switched-memory B lymphocytes. J Immunol Res 2014;2014:635108.

33. Fernàndez JE, Deaglio S, Donati D, Beusan IS, Corno F, Aranega A, et al. Analysis of the distribution of human CD38 and of its ligand CD31 in normal tissues. J Biol Regul Homeost Agents 1998;12:81–91.

34. Medina F, Segundo C, Brieva JA. Purification of human tonsil plasma cells: pre-enrichment step by immunomagnetic selection of CD31(+) cells. Cytometry 2000;39:231–4.

35. Ellyard JI, Avery DT, Phan TG, Hare NJ, Hodgkin PD, Tangye SG. Antigen-selected, immunoglobulin-secreting cells persist in human spleen and bone marrow. Blood 2004;103:3805–12.

36. Nagy Á, Munkácsy G, Győrffy B. Pancancer survival analysis of cancer hallmark genes. Sci Rep 2021;11:6047.

37. Giraldo NA, Sanchez-Salas R, Peske JD, Vano Y, Becht E, Petitprez F, et al. The clinical role of the TME in solid cancer. Br J Cancer 2019;120:45–53.

38. Veglia F, Sanseviero E, Gabrilovich DI. Myeloid-derived suppressor cells in the era of increasing myeloid cell diversity. Nat Rev Immunol 2021;21:485–98.

39. Kuswanto W, Nolan G, Lu G. Highly multiplexed spatial profiling with CODEX: bioinformatic analysis and application in human disease. Semin Immunopathol 2022:1–13.

40. Hickey JW, Tan Y, Nolan GP, Goltsev Y. Strategies for accurate cell type identification in CODEX multiplexed imaging data. Front Immunol 2021;12: 727626.

41. Manrique SZ, Correa MA, Hoelzinger DB, Dominguez AL, Mirza N, Lin HH, et al. Foxp3-positive macrophages display immunosuppressive properties and promote tumor growth. J Exp Med 2011;208:1485–99.

42. Devaud C, Yong CS, John LB, Westwood JA, Duong CP, House CM, et al. Foxp3 expression in macrophages associated with RENCA tumors in mice. PLoS One 2014;9:e108670.

43. Allard B, Panariti A, Pernet E, Downey J, Ano S, Dembele M, et al. Tolerogenic signaling of alveolar macrophages induces lung adaptation to oxidative injury. J Allergy Clin Immunol 2019;144:945–61.

44. de Miranda MC, Melo MIAd, Cunha PdS, Gentilini J, Faria JAQA, Rodrigues MA, et al. Roles of mesenchymal stromal cells in the head and neck cancer microenvironment. Biomed Pharmacother 2021;144:112269.

45. Yoshikawa K, Ishida M, Yanai H, Tsuta K, Sekimoto M, Sugie T. Prognostic significance of PD-L1-positive cancer-associated fibroblasts in patients with triple-negative breast cancer. BMC Cancer 2021;21:239.

46. Joshi RS, Kanugula SS, Sudhir S, Pereira MP, Jain S, Aghi MK. The role of cancer-associated fibroblasts in tumor progression. Cancers (Basel) 2021;13:1399.

47. Zhao H, Wang H, Zhou Q, Ren X. Insights into tertiary lymphoid structures in the solid tumor microenvironment: anti-tumor mechanism, functional regulation, and immunotherapeutic strategies. Cancer Biol Med 2021;18:981–91.

48. Sofopoulos M, Fortis SP, Vaxevanis CK, Sotiriadou NN, Arnogiannaki N, Ardavanis A, et al. The prognostic significance of peritumoral tertiary lymphoid structures in breast cancer. Cancer Immunol Immunother 2019;68:1733–45.

49. Munoz-Erazo L, Rhodes JL, Marion VC, Kemp RA. Tertiary lymphoid structures in cancer – considerations for patient prognosis. Cell Mol Immunol 2020;17: 570–5.

50. Martinez-Murillo P, Pramanik L, Sundling C, Hultenby K, Wretenberg P, Spångberg M, et al. CD138 and CD31 double-positive cells comprise the functional antibody-secreting plasma cell compartment in primate bone marrow. Front Immunol 2016;7:242.

51. Brynjolfsson SF, Persson Berg L, Olsen Ekerhult T, Rimkute I, Wick M-J, Mårtensson I-L, et al. Long-lived plasma cells in mice and men. Front Immunol 2018;9:2673.