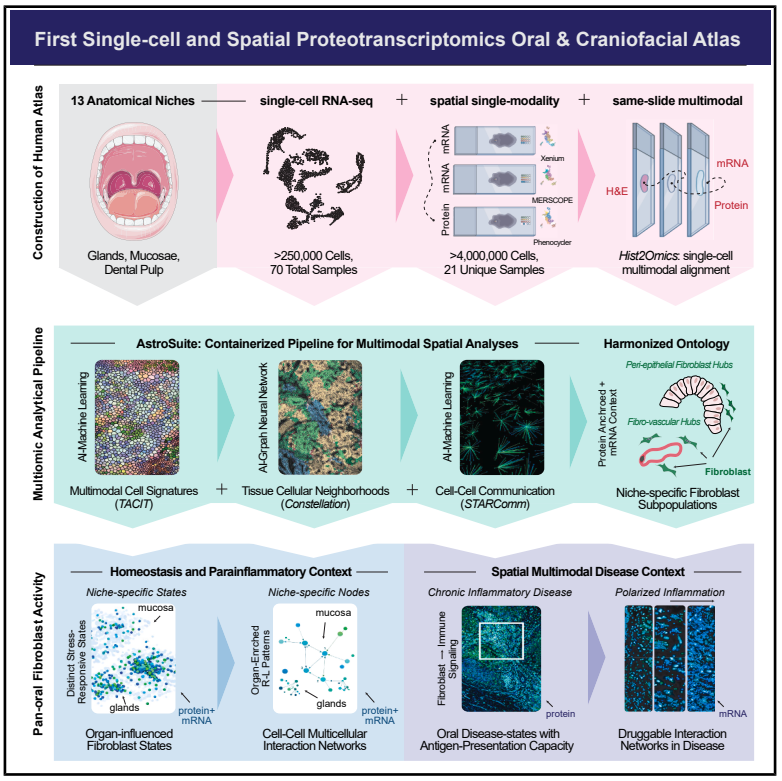
Cell Press Blue

Article

An integrated single-cell and spatial proteotranscriptomics atlas of fibroblast-driven immunoregulation within the human adult oral cavity

Graphical abstract



Highlights

- Integrated single-cell and spatial atlas maps human oral immunoregulatory architecture
- Fibroblasts emerge as central organizers of stromal-immune communication niches
- Spatial proteotranscriptomics defines fibroblast ligand-receptor interaction modules
- Disease integration reveals fibroblast rewiring across inflammatory and reparative niches

Authors

Bruno F. Matuck, Khoa L.A. Huynh, Diana Pereira, ..., Inês Sequeira, Jinze Liu, Kevin M. Byrd

Correspondence

i.sequeira@qmul.ac.uk (I.S.), liuj15@vcu.edu (J.L.), kevinmbyrd@gmail.com (K.M.B.)

In brief

Matuck et al. present an integrated single-cell and spatial proteotranscriptomic atlas of human oral tissues, revealing fibroblasts as central organizers of structural immunity and defining spatial multicellular interaction modules that are rewired in chronic inflammation.

Matuck et al., 2026, Cell Press Blue *1*, 100007
April 20, 2026
https://doi.org/10.1016/j.cpblue.2026.100007

CellPress

Cell Press Blue

CellPress
OPEN ACCESS

Article

An integrated single-cell and spatial proteotranscriptomics atlas of fibroblast-driven immunoregulation within the human adult oral cavity

Bruno F. Matuck,[1,21] Khoa L.A. Huynh,[2,21] Diana Pereira,[3,21] Quinn T. Easter,[1] XiuYu Zhang,[2] Meik Kunz,[4] Nikhil Kumar,[5] Aditya Pratapa,[6] Brittany T. Rupp,[1] Ameer Ghodke,[7] Alexander V. Predeus,[8] Alexandre Fernandes,[3] Lili Szabó,[9,10] Stefan Hartmann,[11] Nadja Harnischfeger,[9,10] Zohreh Khavandgar,[12,13] Margaret Beach,[12,13] Paola Perez,[12] Benedikt Nilges,[14] Maria M. Moreno,[14] Kang I. Ko,[15] Rohit Singh,[6] Purushothama Rao Tata,[6] Sarah A. Teichmann,[16,17] Adam Kimple,[7,18] Sarah Pringle,[19] Kai Kretzschmar,[9,10] Blake M. Warner,[12,13] Inês Sequeira,[3,22,*] Jinze Liu,[2,20,22,*] and Kevin M. Byrd[1,5,20,22,23,*]

[1]Department of Oral and Craniofacial Molecular Biology, Philips Institute for Oral Health Research, Virginia Commonwealth University, Richmond, VA, USA
[2]Department of Biostatistics, Virginia Commonwealth University, Richmond, VA, USA
[3]Center for Oral Immunobiology and Regenerative Medicine, Barts Centre for Squamous Cancer, Institute of Dentistry, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
[4]The Bioinformatics CRO, Sanford, FL, USA
[5]Division of Oral and Craniofacial Health Sciences, Adams School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[6]Department of Cell Biology, Duke University, Durham, NC, USA
[7]Department of Otolaryngology-Head and Neck Surgery, University of North Carolina School of Medicine, Chapel Hill, NC, USA
[8]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK
[9]Mildred-Scheel Early Career Centre (MSNZ) for Cancer Research, University Hospital Wuerzburg, Wuerzburg, Germany
[10]Department of Biochemistry and Molecular Biology, Biocenter, University of Wuerzburg, Wuerzburg, Germany
[11]Department of Oral and Maxillofacial Plastic Surgery, University Hospital Würzburg, Würzburg, Germany
[12]Sjögren's Clinical Investigations Team, National Institutes of Dental and Craniofacial Research, Bethesda, MD, USA
[13]Salivary Disorders Unit, National Institutes of Dental and Craniofacial Research, Bethesda, MD, USA
[14]OMAPiX, Inc, Leuven, Belgium
[15]Department of Periodontics, School of Dental Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[16]Cambridge Stem Cell Institute, Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, University of Cambridge, Cambridge, UK
[17]Department of Medicine, University of Cambridge, Cambridge, UK
[18]Marsico Lung Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[19]University of Groningen and University Medical Center, Groningen, the Netherlands
[20]VCU Massey Comprehensive Cancer Center, Bioinformatics Shared Resource Core, Virginia Commonwealth University, Richmond, VA, USA
[21]These authors contributed equally
[22]Senior author
[23]Lead contact
*Correspondence: i.sequeira@qmul.ac.uk (I.S.), liuj15@vcu.edu (J.L.), kevinmbyrd@gmail.com (K.M.B.)
https://doi.org/10.1016/j.cpblue.2026.100007

SUMMARY

The immunoregulatory architecture of human oral tissues remains poorly defined. We present an integrated single-cell and spatial proteotranscriptomic atlas profiling >250,000 single-cell transcriptomes and >4 million spatially resolved cells across 13 niches. Using our AI-enabled AstroSuite, we defined neighborhoods and interaction modules, revealing peri-epithelial fibroblast-centered hubs enriched in effector cytokines. We harmonized fibroblast subtypes (universal, immune, peri-epithelial, peri-vascular, peri-neural, antigen-presenting cell [APC]-like, stress responsive, and myofibroblasts) with stress-responsive subtypes partitioning between mucosae and glands (type I and II). Spatial multiomics mapped ligand-receptor programs and identified mucosal stress-responsive fibroblasts as putative immunoregulatory hubs. Niche-aware integration of healthy and diseased datasets revealed fibroblast rewiring into inflammatory and reparative niches. Disease neighborhoods exhibited expansion of major histocompatibility complex (MHC)-I^{+}, MHC-II^{+}, and programmed cell death ligand 1 (PD-L1)$^{+}$ fibroblasts and predicted spatial engagement with T cells at tertiary lymphoid structures. Together, this atlas identifies fibroblasts as central regulators of structural immunity and provides a scalable framework to target stromal-immune interactions across barrier organs.

INTRODUCTION

The oral cavity is both barrier and sentinel,[1] serving as a dynamic, immunologically rich interface where the body balances constant environmental exposure with robust internal defense mechanisms.[2,3] As a "window to the body,"[4,5] it contains multiple tissue types, including masticatory, lining and specialized oral mucosae, major and minor salivary glands, and teeth.[6] Despite regional differences in morphology,[7] mucosae share a core architecture of stratified squamous epithelium supported by keratinocytes and underlying stroma populated by fibroblasts, neurovascular units, muscle, and diverse immune cells.[8] Juxtaposed salivary glands, including parotid, submandibular, sublingual, and hundreds of minor salivary glands, are in contact with specific mucosal niches.[9] Despite this proximity, each oral and salivary niche is structurally and functionally distinct,[6] and certain diseases preferentially affect specific sites.[10] However, the cellular and molecular mechanisms underpinning niche-specific disease patterns across these distinct oral tissues remain poorly understood.[11]

Beyond their role as physical barriers, oral mucosae and glands exhibit sensory[12,13] and emerging immunoregulatory functions.[14] Within these tissues, structural cells, such as keratinocytes, fibroblasts, and vascular endothelial cells (VECs), dynamically engage with immune cells.[15,16] Long considered passive tissue-support cells, structural cells are increasingly recognized as active regulators of immune function; this concept of structural immunity acknowledges their capacity to mediate immune responses, maintain homeostasis, and coordinate tissue regeneration.[17,18] Notably, the oral cavity is daily exposed to mechanical, microbial, and dietary stressors that may elicit sub-pathologic inflammation,[19] a form of low-grade, adaptive immune activation, known in other contexts as parainflammation.[20,21] This intermediary state exists between basal homeostasis and classical inflammation, and its molecular programs and spatial organization are not fully characterized. Considering this, fibroblasts, which are strategically positioned across epithelial, peri-vascular, and interstitial zones, appear to be ideal candidates to mediate dynamic inflammatory states.

While recent efforts have mapped fibroblasts across organs and disease states,[22–25] the oral cavity remains largely underrepresented from such atlases. It remains unknown whether fibroblasts orchestrate conserved immune-signaling strategies across distinct mucosae and glands or whether niche-specific subtypes encode unique basal, para-inflammatory, pathologic, and/or immunoregulatory roles. Equally unclear is how this fibroblast-anchored architecture is remodeled in chronic oral disease states, which often feature niche-specific disruption of immune regulation within fibroblast-rich stromal compartments. To address these gaps, we constructed the first integrated single-cell and spatial proteotranscriptomic atlas of adult oral and craniofacial tissues. This resource includes single-cell RNA sequencing data from >250,000 cells from 70 samples representing 13 distinct tissue niches, including oral mucosae, salivary glands, and dental pulp, compiled using new and previously published datasets from the Human Cell Atlas Oral & Craniofacial Bionetwork.[6] We identified major structural and immune cell types, characterized their niche-specific heterogeneity, and profiled ligand-receptor (L-R) expression patterns, revealing that fibroblasts emerged as among the most dynamic interactive cell types, suggesting a central role in orchestrating tissue-specific immune signaling.

To resolve these interactions *in situ*, we created two complementary spatial proteotranscriptomics atlases and subsequently defined standardized terminology to support spatial atlas creation (Box 1). First, we performed same-slide, spatial proteotranscriptomics using a 40-plex spatial-proteomics panel (PhenoCycler-Fusion) after running a custom 300-plex spatial-transcriptomics panel (Xenium). Second, we applied a similar 40-plex spatial-proteomics panel and a similar custom 300-plex spatial-transcriptomics panel (MERSCOPE) to sequential tissue sections. Both atlases were designed to target known fibroblast-driven L-R pairs and applied to 21 unique human samples, capturing nearly 4 million cells across six distinct oral and craniofacial niches. To analyze these data, we applied our recently published spatial analysis suite, AstroSuite,[26–28] including (1) histology-2-omics (hist2omics; single-cell histological and/or multiomic data alignment), (2) TACIT (rapid, hierarchical cell annotation), (3) Constellation (integrated neighborhood identification), (4) STARComm (L-R network analyses), and (5) Astrograph (spatial data visualization). Throughout this study, inferred cell-cell communication reflects spatial co-organization and recurrent expression of signaling components rather than direct biochemical validation of L-R engagement.

Using these datasets, we resolved eight harmonized fibroblast subtypes across oral tissues, including universal, immune, peri-epithelial, peri-vascular, peri-neural, antigen-presenting cell (APC)-like, stress responsive, and myofibroblasts. Among these, stress-responsive fibroblasts were a prominent population in healthy tissues and further divided into two transcriptionally distinct groups: type I, enriched in salivary glands, and type II, enriched in mucosae. Spatial multi-omics revealed that fibroblasts structure tissue neighborhoods and coordinate L-R signaling, with mucosal stress-responsive fibroblasts emerging as potent immunoregulatory hubs. In chronic periodontitis, niche-aware integration showed rewiring of fibroblast states into interdigitated inflammatory and reparative niches, with expansion of major histocompatibility complex (MHC)-I$^+$, MHC-II$^+$, and programmed cell death ligand 1 (PD-L1)$^+$ fibroblasts near tertiary lymphoid structures. Drug2Cell analysis further identified context-specific therapeutic opportunities targeting these immune interfaces. Collectively, these findings position fibroblasts as central immune architects of the oral cavity and establish a scalable framework for precision targeting of stromal-immune ecosystems in health and chronic disease.

RESULTS

An integrated single-cell transcriptomics atlas of human oral tissues

The diversity of human cell types and states remains incompletely defined and, outside of the tooth and periodontium,[18,29] no integrated atlas of the human oral and craniofacial tissues exists. We compiled 11 publicly available single-cell RNA sequencing (scRNA-seq) datasets, including samples from across oral mucosa,[5,30–34] salivary glands,[5,34–37] and dental

Box 1. Key terms

AstroSuite—a modular collection of spatial bioinformatics tools (TACIT, Constellation, STARComm, and hist2omics) enabling integrated identification and visualization of cell types, features, and interactions across multimodal datasets.

Constellation—one of the components of the AstroSuite pipeline that defines tissue cellular neighborhoods (TCNs) by integrating all samples in an experiment and partitioning tissues into spatially co-occurring microenvironments enriched for coordinated cell activity.

hist2omics—one of the components of the AstroSuite pipeline for overlay and analysis of single-cell-resolved spatial omics data from the same slide with histology staining, here applied to spatial proteotranscriptomics combining spatial proteomics, spatial transcriptomics, and H&E.

Multicellular interaction modules (MCIMs)—spatially recurrent multicellular signaling networks identified from ligand-receptor interactions among neighboring cells. They represent coordinated communication programs involving multiple cell types within tissue microenvironments.

Parainflammation—a low-grade, persistent activation of the immune system in tissues, positioned between homeostasis and classical inflammation.

Proteotranscriptomics—spatial measurement of both protein and RNA expression from the same tissue sections at single-cell resolution.

Receptor-ligand density scoring—a STARComm method quantifying spatial density and strength of ligand-receptor (L-R) interactions in integrated tissue datasets.

STARComm—a computational framework quantifying L-R communication and identifying MCIMs based on communication density, spatial expression density, and proximity.

Structural immunity—immune regulation by structural cells (e.g., fibroblasts, endothelial cells) through spatial organization and signaling.

TACIT—an AI-enabled, automated whole-slide classification tool assigning cell types and tissue features from multiplexed spatial imaging data.

Tiered cell annotation—a multi-level cell labeling approach beginning with broad categories and refining to specific identities based on marker expression and tissue context.

Tissue cellular neighborhoods (TCNs)—recurring, spatially organized groups of co-occurring cell types predicted to coordinate local biological functions and signaling.

pulp[31,38,39] (Figures 1A and 1B), and we generated three new scRNA-seq datasets from labial mucosa, soft palate mucosa, and parotid glands to expand coverage and representation of mucosal and glandular niches. We harmonized clinical and sample metadata for all the samples (Table S1). We also generated spatial-proteomics and spatial-transcriptomics data using formalin-fixed paraffin embedded (FFPE) samples (Figure 1C) to create two complementary spatial multiomic atlases: (1) spatial proteomics (PhenoCycler-Fusion) with same-section spatial transcriptomics (Xenium) and (2) spatial proteomics with sequential-section spatial transcriptomics (MERSCOPE), enabling single-cell-resolved integration of proteomic and transcriptomic measurements (Figure 1D).[26]

To interpret the integrated scRNA-seq dataset and enable cross-niche comparisons, we next performed systematic cell-type annotation. We adopted a hierarchical classification approach to capture both broad lineage identities and finer niche-specific subtypes while maintaining consistency with established ontologies.[40] Following a tiered structure (Figure 1E), cells were classified into broad categories (tier 0 [T0]) and refined into eight T1 types, which included structural (epithelial, fibroblast, vascular, neural-crest-derived, muscle, mineralized) and immune (myeloid and lymphoid) populations. T1 annotations revealed consistent detection of most cell types across samples.

The atlas encompasses three principal tissue groups: oral mucosae, salivary glands, and dental pulp, each with distinct cell-type compositions. Only healthy adult samples were included. The dataset skewed toward younger, female donors with limited racial and ethnic diversity (Figure 1F). After normalization and dimensionality reduction (Figure 1G), we benchmarked five batch-correction methods (Harmony, Scanorama, FastMNN, scANVI, and Seurat) using scIB metrics for batch mixing, biological conservation, and cluster separation (Figures S1A–S1C).[41] Harmony achieved the best balance (score, 0.82) and was used for all subsequent analyses. After quality control (see STAR Methods), the final atlas comprised 246,102 cells from 70 samples spanning 13 macroniches, publicly available via CELLxGENE (Human Oral and Craniofacial Cell Atlas [HCA]), including subclustered atlases for major niches and cell types (Figures S1E and S1F).

We computed cell-type signatures for each major cell type (Figure S1D). Epithelial cells were overrepresented in salivary gland and tongue samples, neural-crest-derived populations were enriched in dental pulp, and vascular cells were enriched in mucosae (Figure 1H). Fibroblasts were present across samples, although their proportions varied by niche. Uniform manifold approximation and projection (UMAP) was then generated for all glands and mucosae (see data and code availability). T2 cell annotation across each tissue type increased the total number of identified cell types to 24, enhancing cellular resolution and revealing finer cell heterogeneity (Figure 1I), including some niche-specific epithelial types (glands: myoepithelial, ionocytes, acinar and ductal cells. Mucosae: keratinocytes, melanocytes, Merkel and Langerhans cells). These data establish a foundational dataset of shared and niche-specific cell types across oral tissues.

Structural-immune communication axes across oral tissues

Having assembled the atlas, we first hypothesized that structural cells actively contribute to local immunoregulation through L-R signaling and act as key initiators of immunomodulatory cues. To investigate this, we began by mapping the overall architecture of structural-to-immune cell communication. We first examined tier 1 cell-type annotations in the version 1 atlas across tissues (Figures S1E–S1G). The highest-ranked structural-to-immune interactions were dominated by extracellular matrix

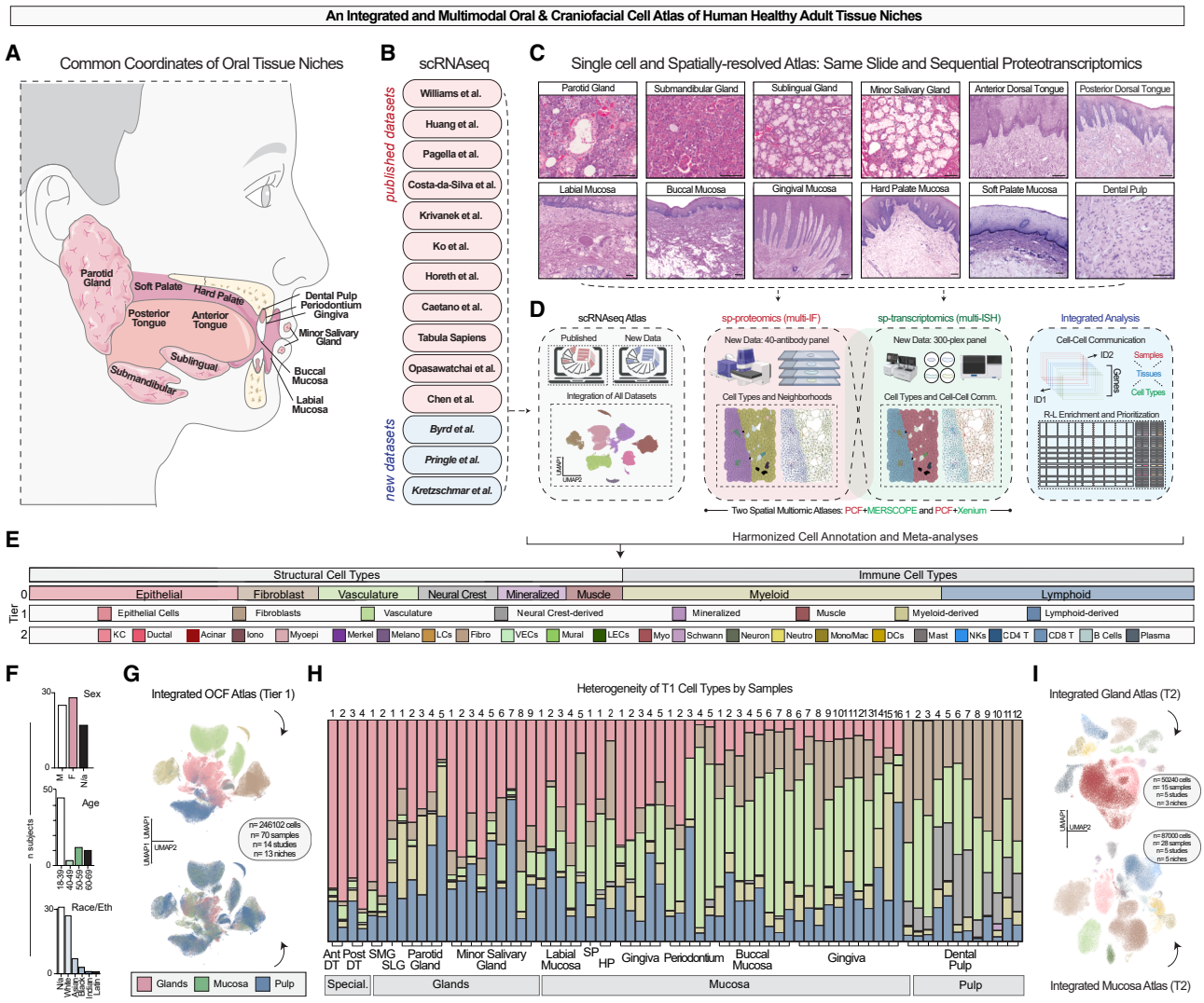


Figure 1. An integrated and multimodal oral and craniofacial cell atlas of human healthy adult tissues

(A) Definition of the CCF of oral and craniofacial macroniches.

(B) Integration of publicly available scRNA-seq datasets (pink) and newly generated datasets (blue).

(C) H&E images of niche-matched formalin-fixed paraffin embedded (FFPE) tissue curated for spatial analysis.

(D) Overview of multimodal integration: the scRNA-seq atlas (gray), the multiplex immunofluorescence (spatial proteomics, red), and the multiplexed *in situ* hybridization atlas (spatial transcriptomics, green). Integrated analyses (blue) combined.

(E) Tiered annotation framework.

(F) Demographics of donors in the scRNA-seq atlas.

(G) UMAP plots of the Harmony integrated tier 1 oral and craniofacial (OCF) atlas (n = 246,102 cells from 70 samples, 14 studies, 13 niches), colored by cell type (top) and tissue niche (bottom). Glands, pink; mucosa, green; pulp, blue.

(H) Distribution of tier 1 cell types across donor samples by niche.

(I) Tier 2 subclustering of integrated gland and mucosal datasets, carried over from the integrated atlas in (G), with colors matched to tier 1 categories in (E).

Iono, ionocytes; KC, keratinocyte; Myoepi, myoepithelial; Melano, melanocytes; LCs, Langerhans cells; Fibro, fibroblasts; VECs, vascular endothelial cells; Myo, myocytes; LECs, lymphatic endothelial cells; Neutro, neutrophils; Mono/mac, monocyte/macrophage; DCs, dendritic cells; NKs, natural killer cells; *T1*, tier one; Special, specialized. Scale bars: glands, 100 μm; mucosal, 250 μm; pulp, 50 μm (C).

(ECM)-derived signals, such as *DCN*, *LUM*, *COL1A2/3A1*, *SPARC*, *PTN*, and *MFAP2* (Figure S1H). We next compared tier 1 cell types between oral mucosa and salivary glands, which differ in function, to determine whether their structural-immune interaction profiles would also vary.

Adapting MultiNicheNet to eight glandular and mucosal niches (see STAR Methods), we modeled L-R signaling across samples and niches simultaneously.[42] These scRNA-seq-based L-R analyses were performed at single-cell resolution using pooled datasets from multiple patients within each tissue group, enabling the

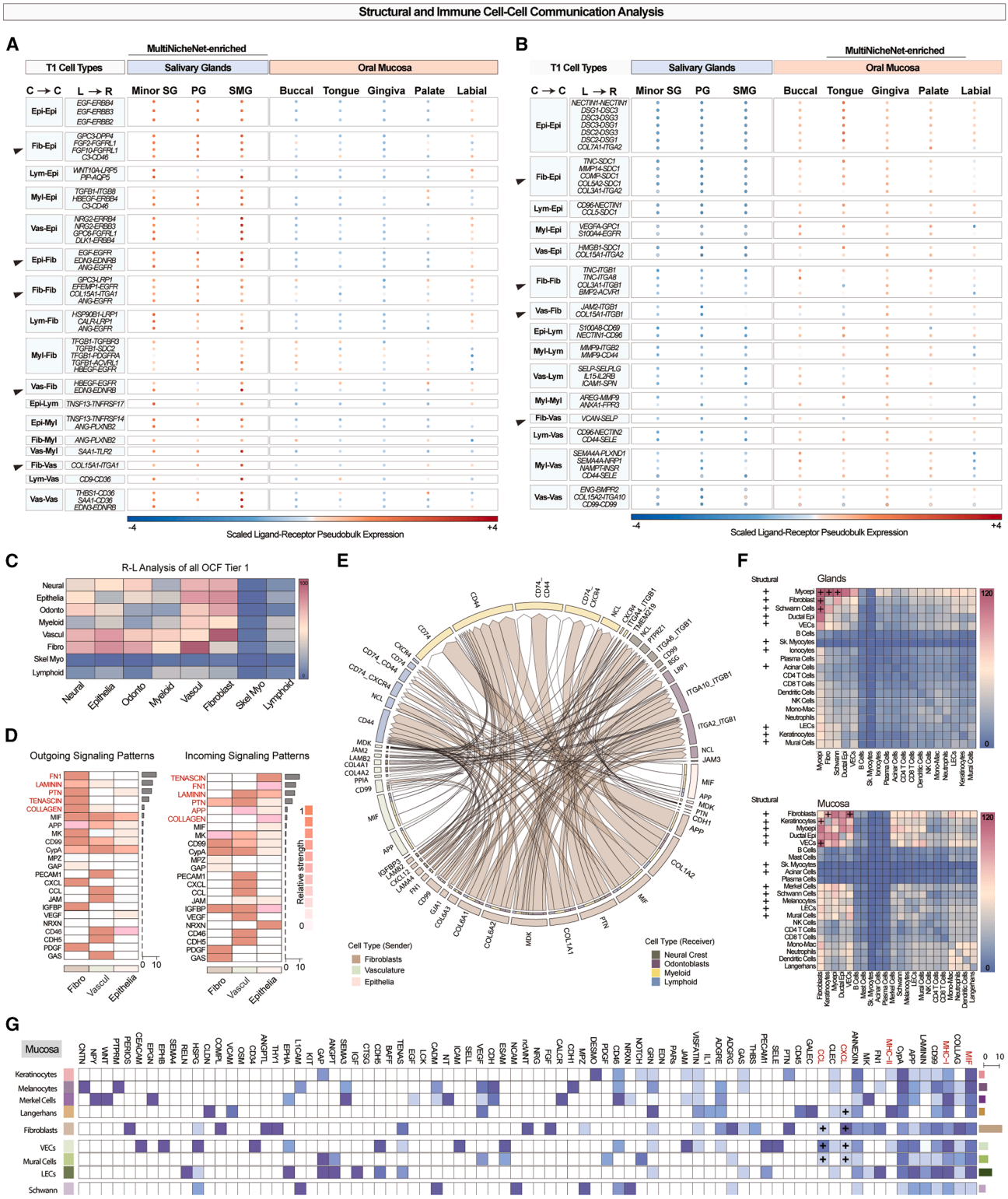


Figure 2. Structural and immune cell-cell communication across salivary glands and oral mucosa

(A and B) MultiNicheNet-enriched ligand-receptor (L-R) analyses showing the top differentially regulated pairs between salivary glands and oral mucosae. (A) Communication enriched in glands relative to mucosa; (B) communication enriched in mucosa relative to glands. L-R interactions were inferred using

(legend continued on next page)

identification of communication programs that are recurrent across samples rather than driven by individual donors. We again identified a prominent fibrovascular signaling axis that varied between tissue types (Figures 2A and 2B; Table S2). Across niches, fibroblasts exhibited extensive but diverse ECM-associated ligand expression programs, underscoring their known roles as key architects of tissue organization. For example, in salivary glands, fibroblast-epithelial communication was anchored by *FGF2-FGFRL1* and *FGF10-FGFRL1*, pathways central to epithelial growth and ductal maintenance. Reciprocal epithelial fibroblast signals, including *EGF-EGFR* and *ANG-EGFR*, reinforced bidirectional signaling circuits. Fibroblast-VEC L-R expression pattern, such as *HBEGF-EGFR*, underscored fibrovascular crosstalk with potential roles in angiogenesis (Figure S2A). In oral mucosae, fibroblast-epithelial signaling was dominated by *TNC-SDC1* and *COL3A1-ITGA1*. Fibroblast-fibroblast modules, including *TNC-ITGB1*, vascular-fibroblast (*JAM2-ITGB1*), and fibroblast-vascular (*VCAN-SELP*) interactions highlighted bidirectional activity that could be related to immune cell trafficking.

Using CellPhoneDB, we quantified overall L-R activity, focusing on these key structural cell types within the integrated atlas.[43] Fibroblasts and VECs displayed the highest levels of interaction potential, followed by epithelial and neural populations, with immune cells acting predominantly as receivers (Figure 2C; Table S3). Pathway-level analyses of outgoing (ligand) and incoming (receptor) signals supported these trends (Figure 2D). Fibroblasts were especially enriched in IGFBP- and CXC-motif chemokine pathways, while vasculature utilized CC-motif chemokines. Using CellChat, we identified shared signaling circuits among epithelia, fibroblasts, and vasculature, often partitioning between ECM and immunomodulatory signaling (STAR Methods; Figure 2E).

To capture structural-immune interaction diversity, we reanalyzed data with refined T2 cell-type resolution, revealing enriched networks in oral mucosae (Figure S2B). L-R patterns across glands and mucosa again showed fibroblasts and VECs as major hubs, ranking top 5 in both (Figure 2F; Table S3). In glands, fibroblasts signaled with epithelial cells, suggesting roles in architecture and homeostasis. In mucosa, fibroblasts interacted with keratinocytes and myeloid cells via chemokine pathways (Figure 2G), including CCL2-CCR2 (monocyte/DC recruitment), CCL5-CCR1 and CXCL12-CXCR4 (mast/natural killer [NK] interactions), and CCL19-CCR7 (lymphocyte homing). Fibroblasts were also predicted to engage MHC-I and CD99-mediated pathways (Figure S2C), linking architecture to immunoregulation.

Fibrovascular-anchored neighborhoods coordinate divergent immune states among niches

We reasoned that the spatial organization of oral mucosa and salivary glands not only reflects tissue-specific structure but also actively shapes and sustains immune activity through coordinated interactions between structural and immune compartments. Informed by our T2 scRNA-seq data, we generated a high-resolution spatial-proteomic atlas using a custom-designed panel of 40 oligo-conjugated antibodies applied to six healthy oral tissue types: three mucosal niches (tongue, gingiva, buccal) and three glandular (parotid, submandibular, minor; Figures 3A and S3A–S3C). Tissues were selected to have comparable demographics across groups (Table S3).

The antibody panel was validated across disease and health contexts using the Akoya PhenoCycler-Fusion 2.0 platform, and organ-mapping antibody panels (OMAPs) for these six tissues were included as part of the ninth release (v2.3) of the Human Reference Atlas.[44] Whole-slide images were segmented using Cellpose3 with H&E-guided, human-in-the-loop training (Figure S3D).[45] Initial results revealed marked differences in cellular composition and metacellular architecture between mucosal and glandular tissues, suggesting fundamentally distinct structural-immune landscapes (Figures 3B, 3C, S3E, and S3F).

To enable such analyses at scale, we developed a modular and containerized spatial analysis toolkit, AstroSuite,[26–28] which includes TACIT, a machine learning-based classifier for cell types and states; Constellation, a framework for integrating and identifying tissue cellular neighborhoods (TCNs); STARComm, a computational pipeline for mapping L-R signaling into multicellular interaction modules (MCIMs) and organizing them into spatially resolved communication networks; and Astrograph, a tool for spatial data visualization (Figure 3D). Across both spatial-proteomics runs, including 36 total tissues, we identified ~2 million segmented cells that were used to annotate cell types, states, and neighborhoods. As expected, mucosae displayed a higher proportion of innate immune cells (e.g., macrophages, $p = 0.012$), while glands were enriched for adaptive cells (e.g., $CD8^+$ T cells, $p = 0.006$; Figure S3G). Furthermore, using

CellPhoneDB and summarized as scaled pseudobulk expression values. C → C, cell-type sender-receiver relationship; L → R, ligand-receptor. Rows are grouped by structural and immune cell types, with blue indicating lower and red indicating higher scaled expression.

(C) Heatmap of inferred L-R interaction across all T1 cell types.

(D) CellChat-derived analysis of outgoing (ligand) and incoming (receptor) signaling pathways across T1 cell types.

(E) Chord diagram of the top 50 significant L-R interactions across all niches, with edges representing inferred directional communication between sender and receiver cell types.

(F) Heatmaps showing total number of interactions in T2 structural-to-structural and structural-to-immune L-R signaling patterns across glands (top) and mucosa (bottom).

(G) Pathway-level aggregation of inferred L-R interactions within structural cell populations, reporting scaled interaction scores for CXCL, CCL, MHC, and MIF (macrophage migration inhibitory factor) signaling pathways across epithelial, fibroblast, vascular, and related cell types. C-C denotes sender-to-receiver cell-type relationships; L-R denotes ligand-to-receptor pairs. Rows are grouped by structural and immune cell types, with blue and red indicating lower and higher scaled interaction scores, respectively.

SG, salivary gland; Minor SG, minor salivary gland; PG, parotid gland; SMG, submandibular gland; Epith, epithelial; Fib, fibroblast; Vas, vascular; Lym, lymphoid; My, myeloid; VECs, vascular endothelial cells; LECs, lymphatic endothelial cells. Scale bars: insets, 50 μm (A–D). Statistical testing was performed using one-way ANOVA and two-sided *t* tests.

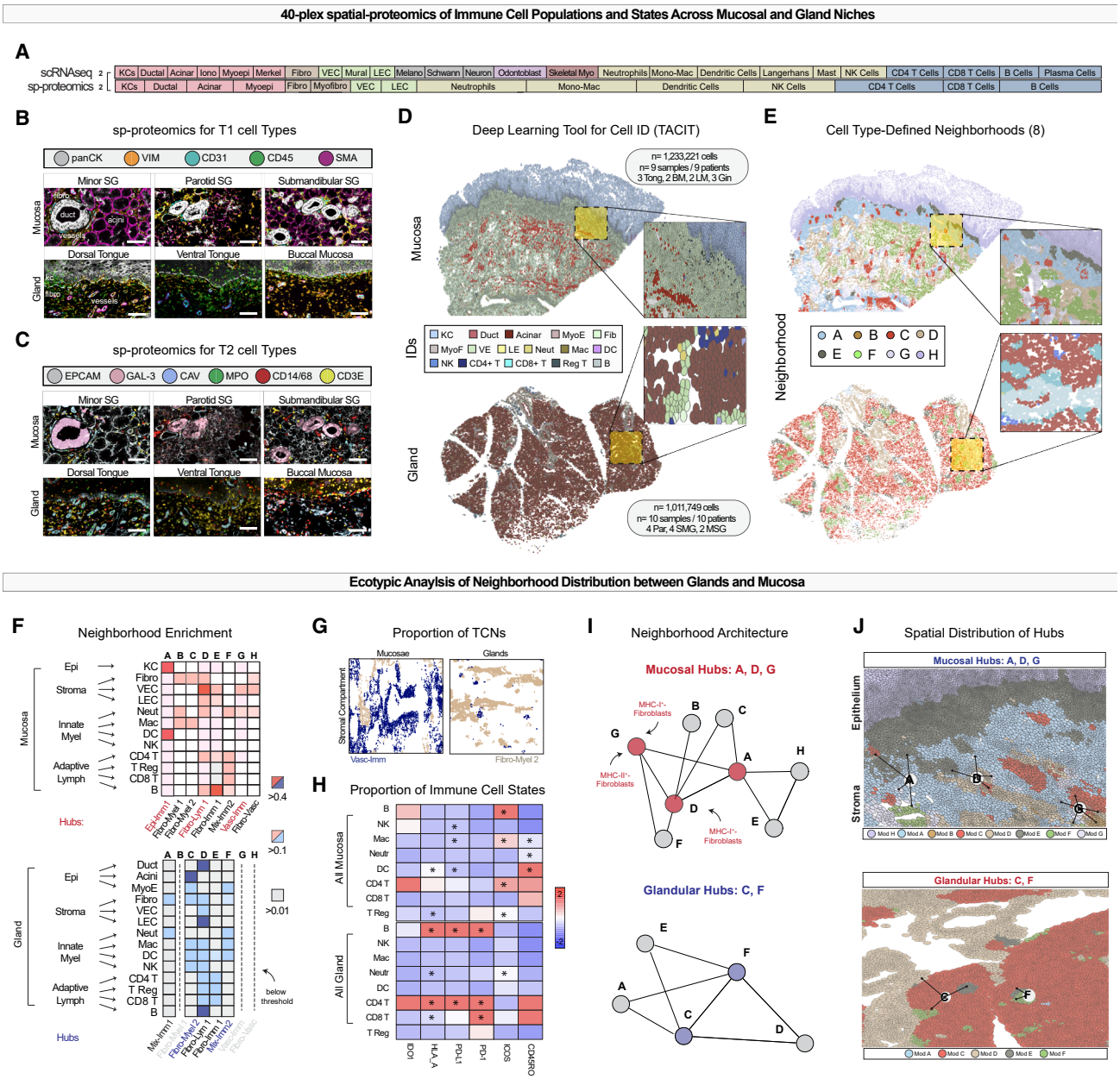


Figure 3. Spatial-proteomics mapping of immune and stromal neighborhoods across oral mucosa and salivary glands

(A) Schematic of tier 1 and tier 2 cell-type classifications.

(B and C) Representative multiplex immunofluorescence images showing tier 1 (B) and tier 2 (C) annotations in mucosa and glands. Cell types were annotated using TACIT (AstroSuite; see STAR Methods).

(D) Deep-learning-based single-cell classification with TACIT.

(E) Tissue cellular neighborhoods (TCNs) identified after whole-experiment integration by Constellation, assigning eight distinct neighborhoods.

(F) Neighborhood enrichment analysis showing epithelial, stromal, and immune composition within TCNs in mucosa (top) and glands (bottom). Neighborhood hubs, which are defined as TCNs with interactions to ≥2 other neighborhoods, are indicated.

(G) Quantification of TCN proportions across mucosa and glands.

(H) Distribution of immune cell subtypes across TCNs stratified by tissue niche.

(I) Network diagrams showing neighborhood architecture and hub connectivity, illustrating fibrovascular hubs in glands (C and F) and mucosa (A, D, and G).

(J) Spatial mapping of mucosal and glandular hubs, with annotated TCNs and surrounding neighborhood structures.

SG, salivary gland; VEC, vascular endothelial cell; TCN, tissue cellular neighborhood; Myofib, myofibroblast; Fibro/Fib, fibroblast; Mac, macrophage; DC, dendritic cell; VEC, vascular endothelial cell. Statistical test: *t* test. Scale bars: 50 μm (B and C).

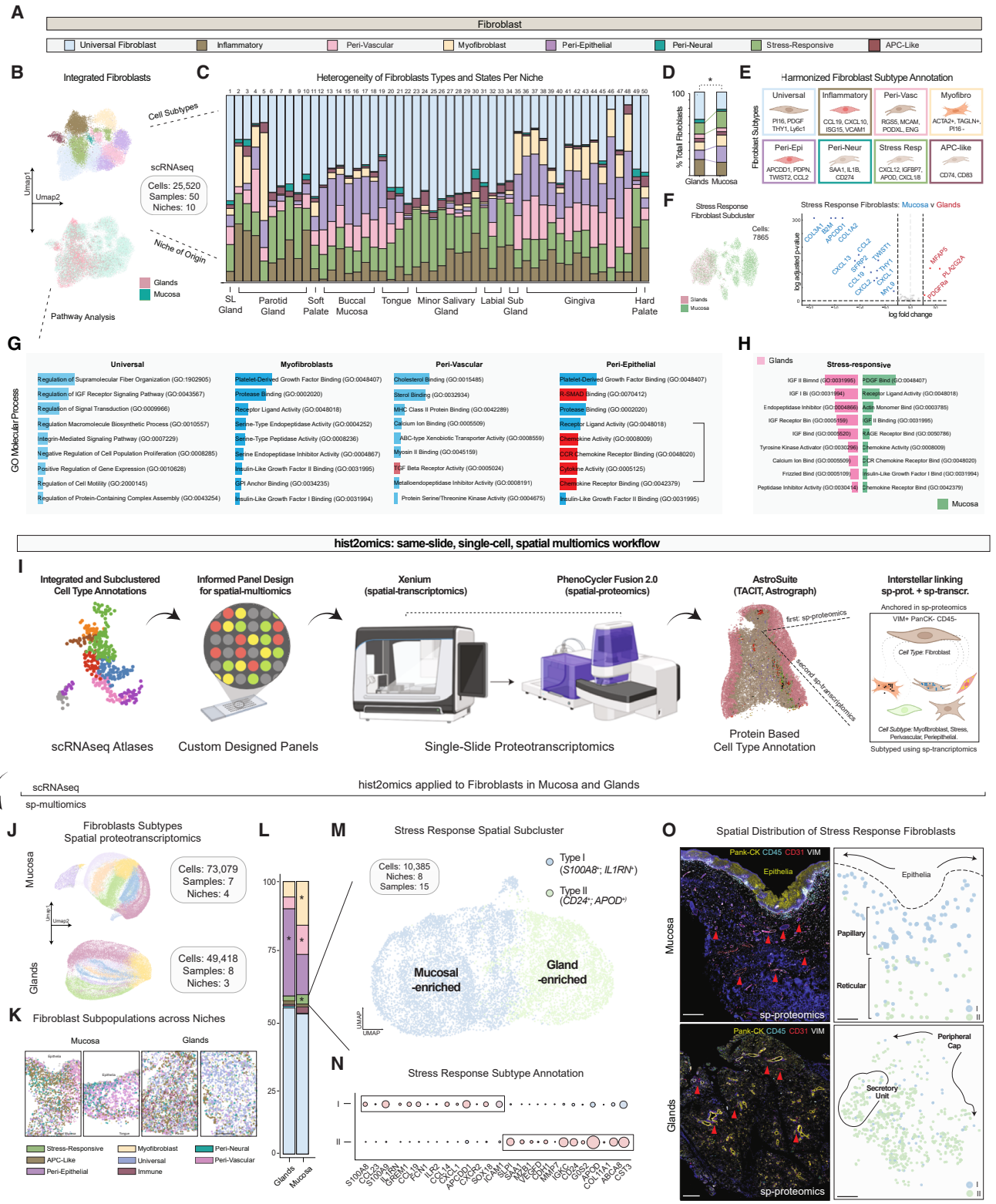


Figure 4. Fibroblast diversity and niche-specific polarization across oral mucosa and glands revealed by scRNA-seq and spatial proteo-transcriptomics

(A) Fibroblast subtype color key, showing eight tier 3 subtypes defined by integrated scRNA-seq.

(B) UMAP projection of >25,000 fibroblasts from 50 samples and 10 oral niches, colored by subtype (top) or tissue of origin (bottom).

(legend continued on next page)

co-occurrence analysis, we found immune cells were significantly closer to structural cells in mucosae than in glands (Figure S3H), indicating tighter immune-structural coupling at barrier surfaces.

To define TCNs, we scanned overlapping tissue regions and integrated patterns into consensus neighborhoods using Constellation (Figure 3F), identifying eight TCNs (A–H) across the dataset. Mucosae exhibited all eight, while glands harbored only five (A and C–F). Fibroblasts were enriched in four mucosal TCNs (B, C, D, and H), with D and H also enriched in VECs, suggesting fibrovascular TCNs resembling "adventitial fibroblasts" from recent atlases. Innate immune cells were broadly distributed across TCNs, while adaptive immune cells preferentially localized to fibrovascular TCNs. Some TCNs were more abundant in mucosa than in glands (Figure 3G), often at peri-epithelial zones, suggesting niches for innate immune surveillance and epithelial barrier maintenance. Divergence of TCNs and associated cell types also showed divergent cell states (Figure 3H).

We next analyzed spatial connectivity between TCNs to identify architectural network organization. We defined TCN "hubs" as neighborhoods connected to $\geq$3 others (Figure 3I). In mucosae, TCN-A, -D, and -G emerged as major hubs, representing mixed-immune, fibro-lymphoid, and vascular-immune neighborhoods, respectively. In glands, TCN-C and -F served as central hubs. While fibrovascular hubs were present in both tissues, those in mucosa were more spatially fragmented, reflecting potential para-inflammatory phenotypes related to structural-immune rewriting of the local niche (Figure 3J). Furthermore, we found higher expression of MHC-I$^+$ (HLA-A$^+$) fibroblasts in mucosal TCN hub D and MHC-I$^+$ and MHC-II$^+$ (HLA-DR$^+$) fibroblasts in mucosal TCN hub G (Figure S3I).

scRNA-seq resolves fibroblast diversity and niche-specific programs across oral tissues

Building on these spatially defined fibrovascular hub phenotypes, we next sought to resolve the underlying fibroblast diversity and transcriptional states using scRNA-seq to determine whether specific subtypes align with the structural-immune architectures observed *in situ*. To achieve this, we first subclustered all fibroblasts from our harmonized scRNA-seq atlas, comprising >25,000 fibroblasts from 10 niches. Using literature-informed annotations, we resolved eight transcriptionally distinct subtypes (Figures 4A–4C and S4A): universal, inflammatory, peri-vascular, myofibroblast, peri-epithelial, peri-neural, stress-responsive, and APC-like fibroblasts, each enriched in different niches with predicted unique functions.[23,24,46–51]

Proportional analysis demonstrated both shared and unique distributions by niche (chi-squared test; $p < 0.001$; Figures 4D and 4E). Importantly, these stress-associated fibroblast programs do not represent entirely novel cell types but, rather, context-dependent reorganization and polarization of conserved fibroblast stress-response states across epithelial surface and glandular niches. Mucosal fibroblasts contained higher proportions of stress-responsive fibroblasts (*CXCL12*$^+$, *IGFBP7*$^+$, *APOD*$^+$, *CXCL1*$^+$, *CXCL8*$^+$), whereas glands were enriched for universal (*P16*$^+$, *PDGF*$^+$, *THY1*$^+$) and inflammatory fibroblasts (*CCL19*$^+$, *CXCL10*$^+$, *ISG15*$^+$, *VCAM1*$^+$). Additional subtypes included peri-epithelial (*APCDD1*$^+$, *PDPN*$^+$, *TWIST2*$^+$, *CCL2*$^+$), myofibroblasts (*ACTA2*$^+$, *TAGLN*$^+$, *PI16*$^-$), APC-like fibroblasts (*CD74*$^+$, *CD83*$^+$), peri-neural fibroblasts (*SAA1*$^+$, *IL1B*$^+$, *CD274*$^+$), and peri-vascular fibroblasts (*ACTA2*$^-$, *RGS5*$^+$, *MCAM*$^+$, *PODXL*$^+$, *ENG*$^+$), each with molecular programs suggestive of specialized roles (Figures 4E and 4F). Stress-responsive fibroblasts comprised ~25% of all fibroblasts, yet subclustering revealed divergent niche associations. Mucosa-enriched stress-responsive fibroblasts upregulated matrix-remodeling genes (*COL1A2*$^+$, *COL3A1*$^+$, *MYL9*$^+$), immune modulators (*CXCL1*$^+$, *CXCL2*$^+$, *CXCL13*$^+$, *CCL2*$^+$, *CCL19*$^+$), and epithelial-mesenchymal regulators (*TWIST1*$^+$, *SFRP2*$^+$, *APCDD1*$^+$).

By niche, we observed differences as well, highlighting the plasticity and dynamic potential of fibroblast populations across oral niches, supporting context-dependent functional specialization within shared structural archetypes. We were interested in the stress-responsive state within both mucosal and gland populations, as they were about 25% of the fibroblasts and, upon subclustered analyses, they displayed divergent clustering by niche (Figure 4F). Notably, mucosa-enriched stress-responsive fibroblasts displayed coordinated upregulation of matrix-remodeling genes (*COL1A2*, *COL3A1*, *MYL9*), immune modulators (*CXCL1*, *CXCL2*, *CXCL13*, *CCL2*, *CCL19*), and epithelial-mesenchymal regulators (*TWIST1*, *SFRP2*, *APCDD1*), suggesting a hybrid state tuned for barrier maintenance, inflammatory

(C) Proportional distribution of fibroblast subtypes per niche.
(D) Overall subtype composition in glands versus mucosa.
(E) Marker gene signatures used for fibroblast subtype annotation.
(F) Differential expression between mucosa and gland-enriched stress-responsive fibroblasts.
(G and H) Gene Ontology (GO) term enrichment analyses for universal, myofibroblast, peri-vascular, and peri-epithelial fibroblasts (G) and stress-responsive fibroblasts stratified by tissue (H). Red color scale highlights immune pathways achieving significance.
(I) Hist2omics workflow: linking spatial transcriptomics and spatial proteomics on the same slide. Cell annotation anchored by vimentin$^+$, Pan-CK$^-$, CD31$^-$, CD45$^-$ cells for transcriptomic deeper subtype annotation.
(J) UMAP projections of spatially annotated fibroblasts from mucosa (73,079 cells, seven samples, four niches) and glands (49,418 cells, eight samples, three niches).
(K) Spatial maps of fibroblast subtype distribution within representative mucosal and glandular niches.
(L) Quantification of subtype abundance across tissues.
(M) Stress-responsive fibroblast subclustering resolves two populations: type I (mucosa enriched) and type II (gland enriched).
(N) Marker gene expression profiles distinguishing type I (S100A8, IL1RN) and type II (CD24, APOD) stress-responsive fibroblasts.
(O) Spatial localization of type I and type II stress-responsive fibroblasts within mucosal and glandular tissues; arrows indicate small vessels.
PCF, PhenoCycler-Fusion; VIM, vimentin; Pan-CK, pan-cytokeratin; VEC, vascular endothelial cells. Statistical test: one-way ANOVA and *t* test. Scale bar: 500 μm.

signaling, and dynamic tissue remodeling as suggested by spatial-proteomics data (Figure 3). Consistent with our scRNA-seq findings, this mucosal stress-response program may underpin heightened immune vigilance and epithelial resilience in barrier tissues, contrasting with the gland-predominant fibroblast states enriched for homeostatic or inflammatory roles.

Pathway enrichment analysis further underscored these niche-specific programs, with universal fibroblasts enriched for supramolecular fiber organization and insulin growth factor (IGF) receptor signaling. Myofibroblasts were enriched for PDGF binding and protease activity, consistent with contractile and matrix-remodeling phenotypes that enable tissue reinforcement during repair. Peri-vascular fibroblasts displayed cholesterol/sterol binding and MHC class II activity, implicating them in lipid trafficking to vascular endothelium and potential perivascular antigen presentation. Peri-epithelial fibroblasts also exhibited PDGF binding together with pronounced chemokine and cytokine activity, positioning them as key modulators of epithelial proliferation and immune cell recruitment at barrier surfaces (Figure 4G).

Focusing on stress-responsive fibroblasts, mucosa-enriched populations showed preferential enrichment for PDGF binding, L-R activity, actin monomer binding, and RAGE/chemokine-receptor binding, all hallmarks of immune surveillance. In contrast, gland-predominant stress-responsive fibroblasts favored IGF-I/II binding, endopeptidase inhibition, tyrosine kinase activation, and calcium binding, suggesting a shift toward metabolically supported reparative programs that preserve secretory architecture and function under stress (Figures 4G and 4H). Taken together, these scRNA-seq findings resolve a fibroblast taxonomy with distinct niche-specific transcriptional programs and predicted functional polarizations, linking structural localization to specialized roles in ECM dynamics, immune modulation, and tissue-specific resilience. These data also suggest that niche stress on fibroblasts can be intrinsic and distinct.

Spatial validation of fibroblast subtypes and niche-dependent polarization

We next tested whether these fibroblast states were spatially organized and preserved *in situ*. To achieve this, we established and applied a new hist2omics framework to same-slide spatial proteotranscriptomic data, using a custom Xenium panel targeting scRNA-seq-informed fibroblast subtype markers (Figure 4I). The proteotranscriptomic datasets were overlaid at single-cell resolution through image-based registration using their respective DAPI-stained nuclear images.[26,52,53] Fibroblast identity was anchored to the spatial-proteomics dataset by selecting vimentin/(VIM)$^+$, CD31$^-$, CD45$^-$, Pan-CK$^-$ cells, thereby excluding endothelial, immune, and epithelial lineages. Segmentation masks generated from spatial proteomics were transferred to the spatial-transcriptomics dataset, enabling deeper annotations restricted to the vimentin$^+$ population.

This protein-first anchoring was essential to harmonizing with the scRNA-seq data, as spatial-transcriptomics annotation alone failed to capture a substantial fraction of confirmed fibroblasts seen by proteomics (Figures S4A–S4C), often pulling in other cell types not intended to be studied here. Once filtered, VIM+, CD31$^-$, CD45$^-$, Pan-CK$^-$ cells displayed similar signatures to scRNA-seq that were overall consistent by niche (Figures S4D and S4E). As expected, the effector cytokine expression was highest in our stress-responsive and inflammatory fibroblasts states as expected (Figure S4F). Considering glands and mucosae, we found that the eight fibroblast types were distributed logically between epithelial and stromal compartments (Figure S4G). Ultimately, ~122,000 fibroblasts were identified robust but exhibited niche-specific heterogeneity (Figure 4J). For example, mucosal peri-epithelial zones were enriched for peri-epithelial, inflammatory, stress responsive, and myofibroblast populations, often forming spatial gradients toward epithelial borders. Glandular fibroblasts favored universal and immune-like phenotypes with more dispersed spatial patterns (Figure 4K). Analyses showed expansion of stress-responsive fibroblasts and myofibroblasts, modest enrichment of immune and peri-vascular fibroblasts in mucosa, and more peri-epithelial fibroblasts in glands, reflecting acinar and ductal dominance (Figure 4L).

We again focused on our defined and divergent stress-responsive fibroblasts, which represented approximately 10% of total fibroblasts in spatial data. Unsupervised subclustering distinctly resolved our stress responsive fibroblasts into two stress-state populations, now termed type I and type II stress-responsive fibroblasts, with divergent niche distributions (Figure 4M). Type I fibroblasts, enriched in mucosa as seen in scRNA-seq data, expressed inflammatory surveillance, epithelial–mesenchymal plasticity, and matrix-remodeling genes (*S100A8/9*, *CCL19*, *IL1RN*, *CXCL1*, *GREM1*, *CXCR2*, *ICAM1*; Figure 4N). Type II fibroblasts, predominant in glands, expressed repair, immune modulatory, and lipid metabolism genes (*CST3*, *ABCA8*, *COL11A1*, *APOD*, *CD24*, *IGKC*, *MMP7*, *VEGFD*, *SAA1*, *SLP1*). Spatial localization confirmed these niche-specific programs as type I fibroblasts localized primarily to peri-epithelial (papillary) layers of oral mucosae lamina propria, whereas type II fibroblasts were enriched in deeper reticular mucosal layers of lamina propria and peri-epithelial regions of glandular tissue (Figure 4O). These single-cell and spatial analyses show that fibroblast subtypes found by scRNA-seq are preserved in tissues and display niche-specific locations and functions.

Spatial proteotranscriptomics predicts fibroblast-driven interaction hubs

Having identified fibroblast subtypes and zones of immune activation, we sought to determine whether fibroblasts were embedded within reproducible, spatially organized networks of cell-cell communication that could functionally define their niche roles. Inspired by frameworks such as TCNs[54] and recently proposed spatial ecotypes,[55] we hypothesized that L-R interactions in oral tissues might also follow patterned architecture, forming clusters of interacting cells defined not just by type or location but by their spatially coordinated and recurrent cell-cell communication activity. To answer this, we applied our STARComm framework to find L-R neighborhoods, what we define as multicellular interaction modules (MCIMs), to two complementary spatial-transcriptomics datasets in parallel (Figure 5A). MCIMs are defined here as spatially recurrent communication patterns inferred from proximity-constrained ligand and receptor expression rather than direct measurements of receptor occupancy,

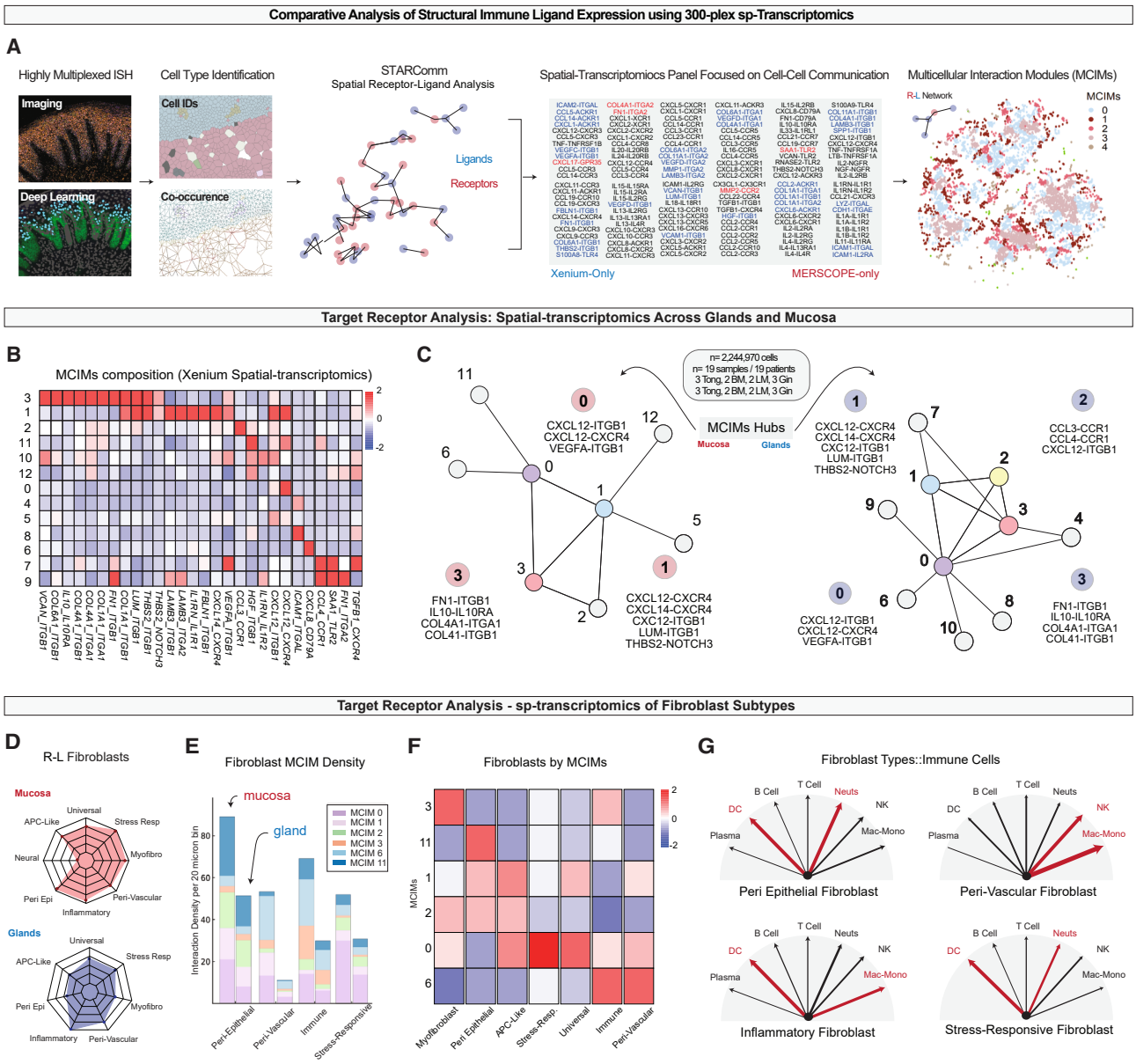


Figure 5. Spatial proteotranscriptomics reveals fibroblast-centered MCIMs across oral mucosae and salivary glands

(A) Overview of the spatial receptor-ligand analysis for cell-cell communication (STARComm) workflow applied to Xenium (∼125 curated L-R pairs) and MERSCOPE (∼100 curated L-R pairs) datasets.

(B) Heatmap showing MCIM composition from Xenium-derived spatial transcriptomics across mucosal and glandular tissues. Rows represent individual MCIMs; columns list representative L-R pairs.

(C) MCIM hubs by tissue type, showing modules with higher inferred interaction scores in mucosa (red) and glands (blue).

(D) Radar plots showing L-R expression profiles of fibroblast subtypes in mucosa (top) and glands (bottom). Arrow thickness represents the CellChat communication score (range 0–1); red arrows denote interactions with $p < 0.05$.

(E) Distribution of fibroblast MCIM density by subtype and niche, showing the relative contribution of each fibroblast population to MCIM clusters.

(F) Heatmap of fibroblast subtype participation across MCIMs, illustrating niche-specific module engagement.

(G) Radar plots showing differences in inferred interaction scores between fibroblast subtypes and immune cell types as computed by STARComm.

ST, spatial transcriptomics; CellChat, cell-cell communication inference framework; p, *p* value. Statistical tests: one-way ANOVA and *t* test.

binding kinetics, or downstream signaling activity. STARComm was applied to spatial datasets comprising samples from multiple individuals analyzed jointly, defining MCIMs based on recurrent, spatially constrained L-R interactions conserved across patients rather than sample-specific signaling events.

In the first dataset, we used spatial transcriptomics (MERSCOPE) with a 300-plex panel (~100 curated L-R pairs, 20% of content) covering mucosae and glands (Table S4), piloting STARComm with a single-modality assay. In the second, we applied hist2omics, a same-slide dual-modality workflow linking Xenium (spatial transcriptomics) with PhenoCycler-Fusion (spatial proteomics) for protein-anchored annotation and high-plex RNA detection. The Xenium panel (~125 curated L-R pairs; Table S5) resolved fibroblast subtypes, with ~25% fibroblast-relevant pairs absent from MERSCOPE. Figure 5A shows Xenium-only (blue) and MERSCOPE-only (red) pairs and illustrates how STARComm integrates *in situ* hybridization, TACIT-based annotation, spatial co-occurrence, and constrained L-R proximity analysis to cluster MCIMs.

MERSCOPE-derived STARComm analysis revealed 17 MCIMs across oral mucosae and glands (Figure S5A). Mucosae supported 15/17 modules (missing modules 2 and 13), while glands contained only 10/17, with some modules (e.g., module 14) uniquely present in minor salivary glands. Several MCIMs were dominated by fibroblast-immune or fibroblast-vascular interactions. Modules 3 and 7, containing *CXCL12-CXCR4*, *CXCL16-CXCR6*, and *CCL14-CCR1* pairs, were highly represented in mucosa and reflected stress-responsive fibroblast (type I) crosstalk with immune partners. Module 5, enriched in *IL10-IL10RA* and *EGF-EGFR*, linked fibroblasts with vascular and epithelial compartments. Fibroblasts were among the most highly signaling populations using a 50-μm radius, with mucosal fibroblast-dominated MCIMs enriched for immune-surveillance and barrier-maintenance pathways. We found niche-specific enrichment of MCIMs (Figure S5B). Across all modules, clustering of L-R pairs (Figure S5C) revealed distinct co-associations, suggesting recurrent signaling programs that span multiple MCIMs. For example, *CXCL12-CXCR4* clustered with several chemokine-chemokine-receptor axes (e.g., *CCL17-CCR4*, *CXCL16-CXCR6*), consistent with coordinated lymphocyte recruitment from vascular networks.

To validate fibroblast-centered MCIMs considering our newly defined cell types and states, we applied STARComm to the proteotranscriptomics dataset (Figure 4). We identified shared and distinct MCIMs across oral mucosae and glands. Across the 13 Xenium-derived MCIMs (Figure 5B), interaction signatures were dominated by two major themes similar to previous L-R analyses: (1) ECM-integrin signaling, with repeated enrichment of collagen (*COL1A1*, *COL4A1*, *COL6A1*), laminin (*LAMA5*, *LAMB3*), and fibronectin (*FN1*) binding to *ITGB1*/*ITGA1*/*ITGA2*, suggesting coordinated stromal-structural anchoring across both niches; and (2) inflammatory chemokine and cytokine axes, including *CXCL12-CXCR4*, *CXCL8-CD79A*, and *CCL4-CCR1* (Figure 5C). The differing number of identified MCIMs (17 in MERSCOPE; 13 in Xenium) reflects the non-overlapping L-R pair used in each dataset, driving finer module resolution and greater total module counts in MERSCOPE.

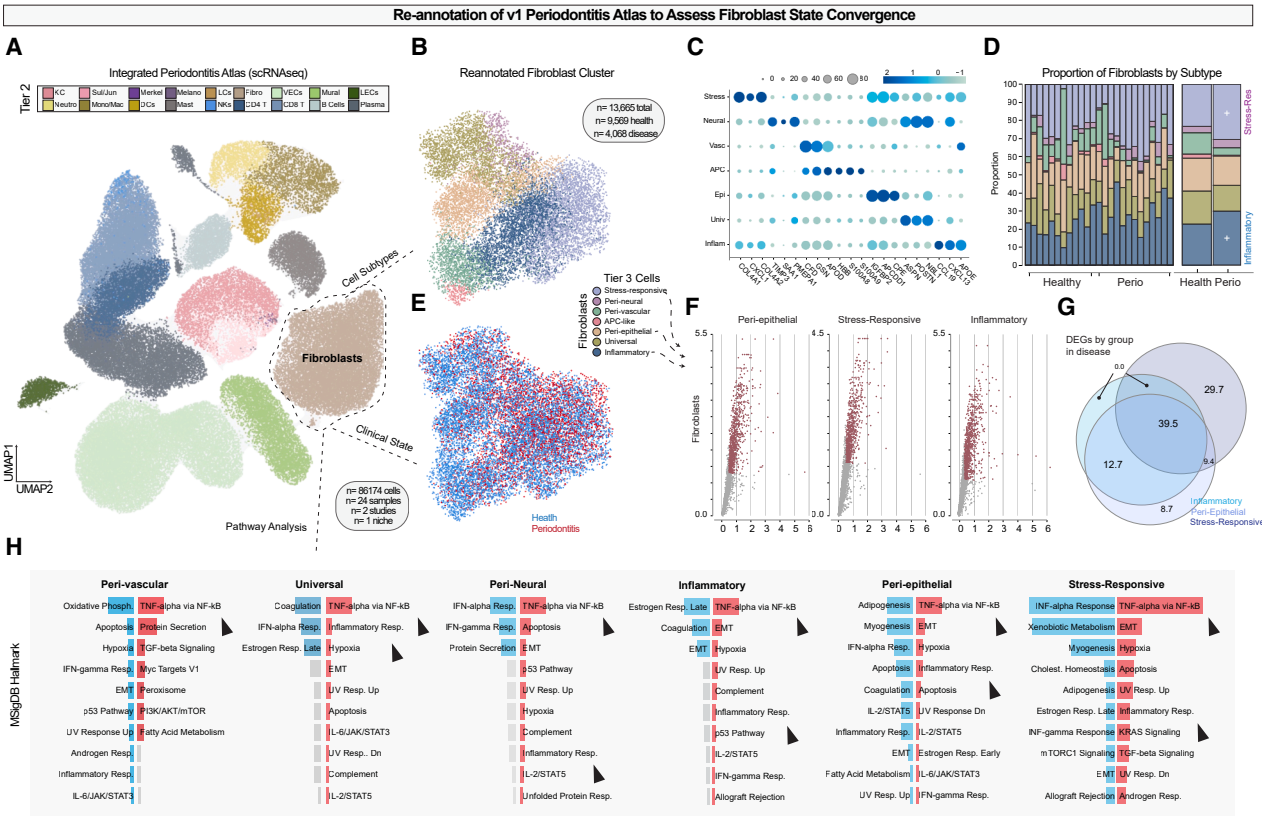
Mucosal tissues showed a broader distribution of modules (MCIMs 0, 1, 3, 6, 11, and 12), whereas glands were enriched for a distinct subset (MCIMs 0, 1, 2, 3, and 4; Figures S5D and S5E). Mucosal MCIMs 1, 11, and 12 showed the highest spatial density, defined by fibroblast-mediated interactions such as *HGF-ITGB1*, *COL4A1-ITGA1*, and *CXCL12-CXCR4*. Gland-associated MCIMs featured a narrower L-R repertoire. MCIM 0 was essentially driven by *CXCL12-CXCR4*, while MCIMs 2 and 4 centered around *CCL3-CCR1*, *ICAM1-ITGAL*, and *VEGFA-ITGB1*. These differences were consistent with prior observations of TCNs, with mucosa and glands maintaining overlapping yet distinct MCIM hubs (e.g., mucosal hubs, MCIMs 0, 1, 3; glandular hubs, MCIMs 0–3). Across both technologies, *CXCL12-CXCR4* emerged as the most conserved and spatially recurrent axis, linking fibroblast-rich niches with immune and vascular guidance across tissue types.

Same-slide proteotranscriptomics linked ligand transcripts to protein-anchored fibroblasts (Figure 4). Overall, ligands were higher in mucosal than in glandular fibroblasts, with inflammatory states highest in both niches (Figure 5D). In mucosa, type I stress-responsive, myofibroblast, and peri-epithelial subsets were also ligand rich, unlike glandular fibroblasts. Mucosal peri-neural fibroblasts showed low L-R activity, likely due to panel limits. In glands, inflammatory fibroblasts dominated, with low counts in type II stress-responsive cells.

Across the major signaling modules (0, 1, 2, 3, 6, and 11), these four shared (peri-epithelial, peri-vascular, immune, and stress responsive) higher enrichments of module 0 (*CXCL12-CXCR4*) and module 6 (*CXCL8-CD79A*) (Figure 5E). When assessing fibroblast subtypes for confirmed L-R pair signatures (Figure 5F), myofibroblasts were enriched for module 3 (ECM ligands across COL, functional neighborhood [FN], VCAN); peri-epithelial fibroblasts for module 11 (*HGF-ITGB1*); and APC-like fibroblasts for *LAMININ*, *IL1RN*, and *CXCL12*. Stress-responsive and universal fibroblasts shared signatures with inflammatory and peri-vascular fibroblasts (module 6). Biologically, inflammatory and peri-vascular fibroblasts preferentially signal to monocytes/macrophages and NK cells. The highest L-R linked axis for T cells was centered on peri-vascular (Figure 5G). Peri-epithelial and type I/II stress-responsive fibroblasts skewed toward an innate immune axis (preferential dendritic cell and neutrophil interactions). By linking fibroblast subtype-resolved hist2omics with STARComm, we mapped local signaling partners for nearly all fibroblast subtypes across innate and adaptive immune populations, suggesting that newly described stress responsive fibroblasts are supportive of innate immune residency within diverse oral tissue niches.

Fibroblast-led immunoregulation and spatial network remodeling in chronic periodontitis

We next applied this approach to interrogate fibroblast spatial architecture and immunoregulatory function in chronic inflammatory disease. Furthermore, we wanted to extend our STARComm-based framework into disease contexts. We again focused on fibroblast activity *in situ*, reanalyzing scRNA-seq datasets from our version 1 human periodontitis scRNA-seq atlas[18]; we also expanded our atlas to generate a multimodal map of health and disease in this niche using the same-slide multimodal (Figures 6A, S6, and S7A). To preserve spatial fidelity, tissue sections were consistently oriented along the tooth-epithelial



(legend on next page)

interface. We hypothesized that chronic inflammatory pressure drives fibroblasts toward spatially anchored immunomodulatory phenotypes that engage and regulate innate immune and lymphocyte-rich immune compartments.

Applying our updated fibroblast annotation schema, we re-clustered fibroblasts across 24 samples (n = 13,665 cells), recovering the same T3 subpopulations: stress-responsive, inflammatory, peri-epithelial, peri-neural, peri-vascular, APC-like, and universal fibroblasts (Figures 6B and 6C). In periodontitis, the proportions of stress-responsive and inflammatory fibroblasts appeared to expand, while universal, APC-like, and peri-neural populations contracted (Figure 6D). When comparing genes upregulated in periodontitis compared to controls, most differences were found in inflammatory and stress-responsive clusters (Figures 6E and 6F). Peri-epithelial fibroblasts, which are adjacent to disease-originating niche in periodontitis stage III/grade B, shared substantial transcriptional overlap with these pro-inflammatory clusters, suggesting a plastic injury-response program near site of disease origination (Figure 6G). MSigDB analysis identified strong tumor necrosis factor (TN)F-α, nuclear factor (NF)-κB, and transforming growth factor (TGF)-β pathway activation in type I stress-responsive fibroblasts, the highest among subpopulations (Figure 6H), pointing to an underappreciated role for these new cell states in immunoregulation.

To spatially contextualize this fibroblast ligand expression (Figures S7C and S7D), our MERSCOPE data were reanalyzed with a focus on structural-immune ligands (Figure 6I). We mapped fibroblast activity along the lesion interface, Importantly, in periodontitis, fibroblasts exhibited upregulation of pro-inflammatory mediators such as *TNF*, *IL1B*, *CXCL8*, and *AREG*, with fibroblasts showing the most pronounced signaling increases compared to other structural cell types (lymphatic endothelial cells [LECs], mural cells, VECs; Figure 6J). L-R analyses revealed a marked expansion of fibroblast-directed immune signaling across chemokine, interleukin, and *EGFR*-associated modules (Figure 6K).

Fibroblasts remained the highest-predicted interaction partners, both homotypically (fibroblast-fibroblast) and heterotypically with immune cells (fibroblast-immune; Figures S7E and 6L). Considering tissue-wide rewiring in disease, the number of MCIMs increased from 13 in healthy mucosa to 20 in disease (Figure 6M). Disease-specific MCIMs (e.g., 3 and 7) showed expanded $CD4^+$ and $CD8^+$ T cell interactions. Disease hub modules were enriched for *CCL2-CCR2*, *CX3CL1-CX3CR1*, *AREG-EGFR*, *TNF-TNFRSF1B*, and *IL3-IL3RA*, many of which were fibroblast specific. While individual fibroblast-immune axes have been described in other chronic inflammatory conditions, no prior periodontitis studies have mapped such disease-specific MCIMs linking fibroblasts to $CD4^+$/$CD8^+$ T cells within a spatially resolved multicellular network.

Spatial module analysis in and around tertiary lymphoid structures (TLSs) revealed fibroblast-immune MCIMs concentrated both within and surrounding these immune aggregates (Figure 6N). TLS-enriched modules (1, 7, 8, 13, 14, 17) were dominated by fibroblast *CCL2*, *CCL5*, *CX3CL1*, and *IL1B* signaling, whereas surrounding modules (2–6 and 20) featured broader cytokine and EGFR ligand signaling (*IL6*, *TNF*, *CXCL9/10/12*, *AREG*, *HBEGF*). TLS borders exhibited a patchwork of fibroblast programs that simultaneously recruited immune cells, promoted inflammation, and engaged wound-repair pathways, particularly within peri-epithelial zones (Figure 6O).

To identify therapeutic entry points, we filtered fibroblast-immune L-R pairs enriched in TLS zones and applied our spatial Drug2Cell workflow, revealing that nearly all were targetable with existing compounds (Figures 6P and S7F). Top nominated axes included granulocyte-macrophage colony-stimulating factor (GM-CSF)/CSF2 (sargramostim), IL-6 (satralizumab), and IL-2RA/CD25 (basiliximab), alongside T cell-directed agents such as muromonab-CD3, consistent with fibroblast-linked immunoregulation in TLS-associated MCIMs (Figure S7G). While these druggable axes are not necessarily novel per se, their nomination

Figure 6. Fibroblast-led immunoregulation and spatial network remodeling in chronic periodontitis

(A) Integrated periodontitis atlas (scRNA-seq) re-annotated using updated tier 3 fibroblast state taxonomy.
(B) UMAP of re-clustered fibroblast subtypes (n = 13,665 cells) from healthy and periodontitis gingiva.
(C) Dot plot of fibroblast subtype marker expression.
(D) Proportion of fibroblasts by subtype in healthy vs. periodontitis samples.
(E) Tier 3 cell distribution across health and disease.
(F) Volcano plots showing differential gene expression analysis for peri-epithelial, stress-responsive, and inflammatory fibroblast subtypes between healthy and periodontitis samples.
(G) Venn diagram of shared and unique DEGs among fibroblast subtypes in disease.
(H) MSigDB pathway enrichment analysis results for each fibroblast subtype based on differentially expressed genes. Gray indicates pathways not reaching statistical significance and therefore excluded.
(I) Integrated spatial atlas of gingival mucosa (MERSCOPE 300-plex) with fibroblast-focused ligand expression maps.
(J) STARComm analysis of fibroblast-to-immune L-R interactions between fibroblasts and immune cell types in health and periodontitis.
(K) Heatmap of fibroblast subtype-specific L-R interactions in health and disease.
(L) Network representation of multicellular interaction modules (MCIMs) and co-occurrence relationships in healthy and periodontitis tissues, with labeled L-R pairs contributing to module structure.
(M) Comparison of MCIM composition between health and periodontitis, showing the number of detected modules and the distribution of cell-type interactions.
(N) Spatial mapping of disease-associated MCIMs dominated by fibroblast chemokine and interleukin signaling.
(O) Spatial distribution of fibroblast transcriptional programs associated with immune-related, inflammatory, and tissue-repair gene sets within peri-epithelial regions.
(P) Spatial Drug2Cell analysis of TLS-enriched fibroblast-immune L-R pairs.
APC-like, antigen-presenting cell-like fibroblasts; EGFR, epidermal growth factor receptor; IL, interleukin; NF-κB, nuclear factor kappa-light-chain-enhancer of activated B cells; TGF-β, transforming growth factor β; TLS, tertiary lymphoid structure. Statistical test: one-way ANOVA and t test.

emerges from an *in silico*, spatially resolved framework applied to this specific tissue and disease context, enabling cell- and niche-aware therapeutic inference. Notably, the convergence on agents already evaluated in early-phase clinical trials in related inflammatory oral diseases provides an orthogonal validation of the biological relevance of the predicted interactions. As Drug2Cell only nominates pharmacologically tractable axes by linking expressed targets to known drugs, without distinguishing agonist from antagonist strategies, the biological interpretation must consider context; for example, GM-CSF agonism could exacerbate chronic inflammation in gingiva, whereas antagonism of the same axis could have the opposite effect.[56] Mapping predicted drug sites of action back to tissue confirmed spatially restricted vulnerabilities within fibrovascular hubs.

We then examined our spatial-proteomics data in the context of fibroblast phenotypes to examine changes in TCNs, checkpoint ligand expression, and stromal-immune co-localization in periodontitis. All TCNs identified in health were retained in disease, but their spatial arrangement and cellular composition were markedly altered (Figures 7A and 7B). Additionally, a new VEC/NK cell-enriched TCN was added (TCN-B). The fibrovascular corridor, a structural region enriched in fibroblasts, endothelial cells, and immune populations, fragmented into smaller, spatially discrete TCNs. These localized hubs were found both in peri-epithelial regions adjacent to microbial insult and in peri-osteal zones near the alveolar bone. By immunophenotyping the TCNs, we found significantly upregulated BCL2, CD107A, PD-L1, and ICOS (inducible T-cell costimulator) in TCN-G, a known fibrovascular hub (Figures 7C and 7D). This supports direct fibro-immune interactions were possible, and emergence of this TCN may serve as a biomarker of disease activity.

Interrogation of the scRNA-seq dataset for an expanded immunoregulatory gene panel (Figure S7H) suggested that fibroblasts could express MHC class I/II molecules and multiple immune checkpoint molecules. In APC-like fibroblasts, low-level but co-occurring expression of MHC molecules, PD-L1, and IDO1 was consistent with a capacity for both antigen presentation and immune evasion. To directly assess these features *in situ*, we subclustered fibroblasts from the spatial-proteomics dataset and profiled them for immunophenotypes via multiplex imaging (HLA-A, HLA-DR, BCL2, PD-L1, IDO1, Ki67) (Figures 7E and 7F). This analysis identified nine major fibroblast states defined by immunoregulatory protein expression and/or proliferative activity, including APC/MHC-II$^+$ (HLA-DR$^+$), immune evasive (HLA-A$^+$, PD-L1$^+$), immunoregulatory (HLA-A$^+$, IDO1$^+$), survival primed APC (HLA-DR$^+$, BCL2$^+$), survival primed + immune evasive (HLA-A$^+$, PD-L1$^+$, BCL2$^+$), checkpoint$^+$ (PD-L1$^+$), cycling (Ki67$^+$), mitotic (Ki67$^+$/pHH3$^+$), and immunoregulatory (IDO1$^+$). Fibroblast-associated TCNs in disease were distinguished by upregulation of immunoregulatory molecules such as PD-L1, IDO1, ICOS, CD107a, and the survival factor BCL2 (Figures 7G and 7H).

Both survival-primed APCs and immunoregulatory fibroblasts proportionally increased in disease. PD-L1 was widely expressed across fibroblast-enriched TCNs (Figure 7I). Notably, HLA-DR, HLA-A, and PD-L1 were all increased specifically in fibroblasts in periodontitis, and PD-L1$^+$ fibroblasts showed strong spatial association with programmed cell death protein 1 (PD-1)$^+$ T cells (Figure 7J). While not an exact comparison to our transcriptomics data, some of these fibroblasts are likely some combination of stress-responsive, inflammatory, and APC-like fibroblasts defined by scRNA-seq and spatial-transcriptomics data (Figures 4 and 6). To test whether fibroblasts both expressed immunoregulatory ligands and engaged in spatially anchored signaling with lymphocytes, we applied L-R analysis to the spatial-proteomics dataset, focusing on the MHC-I, MHC-II, and PD-L1-PD-1 checkpoint axis. This followed because an examination of our data predicted many spatially high-confidence PD-L1-PD-1 interactions between fibroblasts (Figure 7K). As observed in the spatial-transcriptomics dataset (Figure 6), transition zones may contain uncharacterized fibroblasts, and MCIM L-R interactions need validation beyond spatial-transcriptomics confirmation of STARComm-inferred hubs (Figure 7L). This provides *in situ* evidence of fibroblast-driven checkpoint engagement with T cells in inflamed gingiva.

We next examined these interactions in relation to TLS-associated niches, focusing on fibroblast-T cell L-R pairs, including HLA-A$^+$ fibroblasts to CD8$^+$ T cells, HLA-DR$^+$ fibroblasts to CD4$^+$ T cells, and PD-L1$^+$ fibroblasts to PD-1$^+$ CD3$^+$ T cells. TLS contained the highest density of these predicted interactions, with PD-L1-PD-1 signaling most frequently linking fibroblasts to regulatory T cells (Tregs), followed by CD8$^+$ T cells, and least to CD4$^+$ T cells (Figure 7M). Given that Tregs are highly migratory and exhibit spatially dependent functional diversity, their prominence in this axis highlights the importance of considering local context when predicting fibroblast-T cell interactions. These results highlight the ability of fibroblasts to adopt multiple immunoregulatory states, positioning them as central modulators of TLS composition and function in chronic periodontitis.

DISCUSSION

This study presents the first integrated single-cell and spatial proteotranscriptomic atlas of adult human oral and craniofacial tissues, ushering in a new era of tissue-resolved mucosal immunology for the oral cavity. Building upon Human Cell Atlas principles, we developed a harmonized framework that unifies cell phenotyping, spatial communication networks, and inflammation states across oral mucosae and glands.[4,54,55] By generating standardized annotations across 13 distinct niches, we reveal how structural cells, particularly fibroblasts, spatially coordinate local immune programs. These findings expand the concept of mucosal structural immunity and highlight fibroblasts as long-lived, spatially embedded regulators of barrier-associated response.

This integrated atlas builds on our prior common coordinate framework (CCF) for the periodontium to establish a pan-oral reference map (Figure 1A).[28,56] The CCF enables consistent identification of structural and immune cells, TCNs, and L-R networks (i.e., MCIMs), forming a scaffold for future spatial analyses and integration into the Human Cell Atlas or the Human BioMolecular Atlas Program (HuBMAP).[46,57] As atlases expand to 2D and 3D reference structures, the ability to anchor observations to anatomical features will be critical for cross-study comparisons and biologic inference.[58,59]

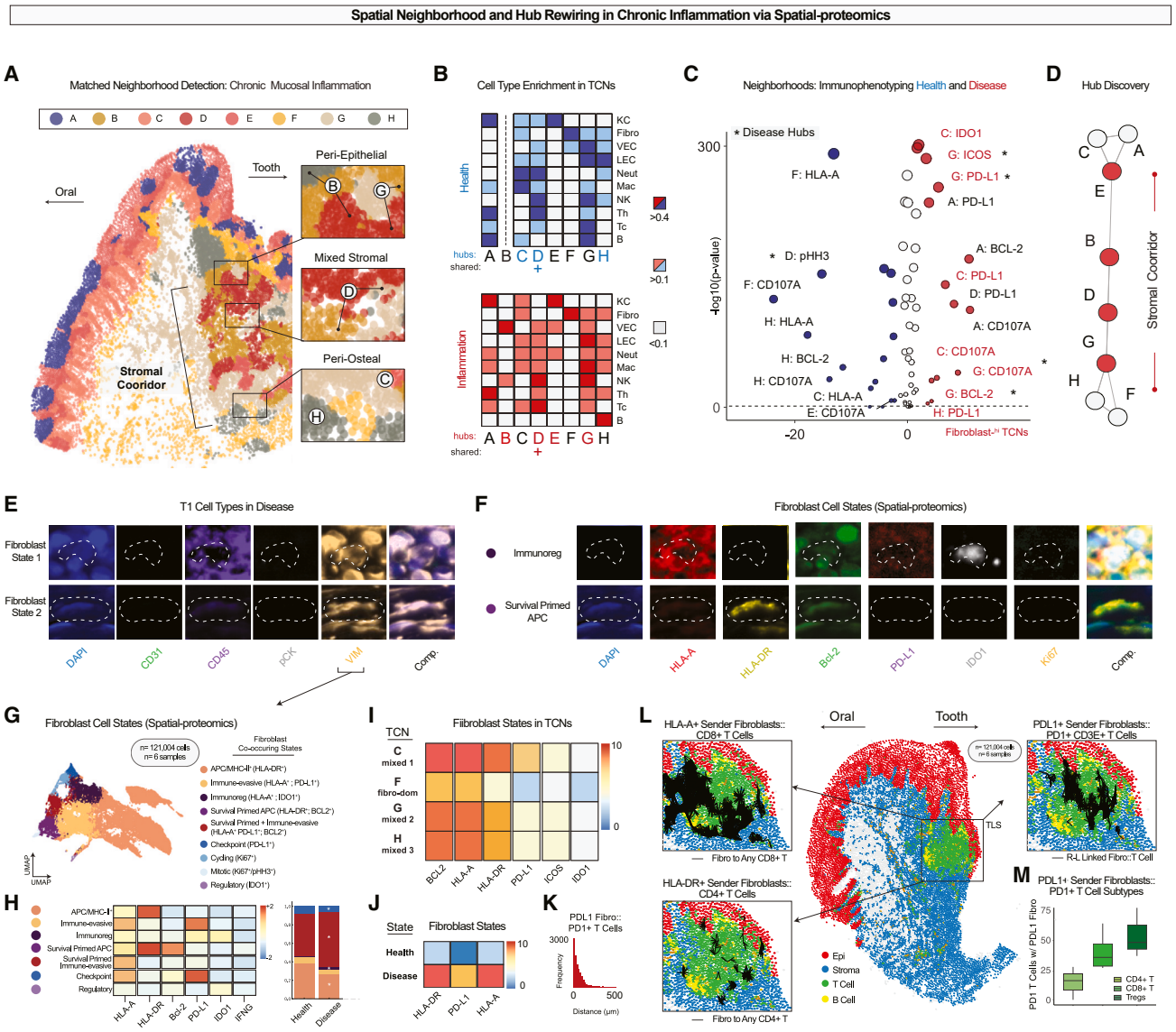


Figure 7. Spatial neighborhood and hub rewiring in chronic inflammation via spatial proteomics

(A) Tissue-wide spatial mapping of matched tissue cellular neighborhoods (TCNs) in chronically inflamed oral mucosa.

(B) Heatmaps showing the distribution of immune (T, B, NK, myeloid) and stromal (fibroblasts, VECs, LECs) cell types across TCNs in health (top) and disease (bottom).

(C) Volcano-plot of immunophenotypic markers across TCNs in specific fibroblast neighborhoods.

(D) Network diagram illustrating disease-driven remodeling of fibrovascular hub connectivity among TCNs.

(E and F) Multiplex immunofluorescence focused on fibroblasts (E) and their states (F) expressing patterns of immunoregulatory molecules.

(G) Spatial map of fibroblast cell states using spatial proteomics across inflamed oral mucosa.

(H) Heatmap of marker combinations across spatially defined fibroblast states and their frequency across health and disease. Stacked bar graph showing proportion of survival primed + immune-evasive (HLA-A$^+$ PD-L1$^+$; BCL2$^+$) as well as immunoregulatory (HLA-A$^+$; IDO1$^+$) expands states in disease.

(I) Distribution of fibroblast states across disease-associated TCNs in inflamed hubs.

(J) Quantification of MHC-I/II and PD-L1 marker expression levels across samples stratified by health and disease status.

(K) Spatial co-localization analysis of PD-L1-expressing fibroblasts and PD-1-expressing CD3$^+$ T cells.

(L) Spatial co-localization of fibroblast-T cell checkpoint interactions. Left top and bottom: fibroblast sender cells (red) and T cell targets (blue) for HLA-A/CD8 and HLA-DR/CD4 axes; right top is an image of PD-L1 (fibroblasts)-PD-1 (CD3$^+$ T cells) predictions.

(M) Histogram showing highest PD-L1-PD-1 pairing in regulatory and cytotoxic CD3$^+$ T cells. Statistical test: one-way ANOVA and *t* test.

A major insight from this study is the spatial anchoring of immune signaling within fibroblast-defined neighborhoods. Across mucosae and glands, fibroblasts emerged as the most transcriptionally interactive structural cell type, exhibiting high ligand expression, spatial connectivity, and anatomical specialization. Using integrated single-cell and spatial proteotranscriptomics, we defined eight harmonized fibroblast subtypes: universal, peri-epithelial, peri-neural, peri-vascular, myofibroblasts, APC-like, immune-regulatory, and two (type I and II) stress-responsive subtypes. These annotations build upon and extend fibroblast ontologies from recent integrated multi-organ atlases[21,22,24] and by incorporating oral tissue relevant transcriptional programs and spatial localization.

Several subtypes show strong parallels with fibroblasts identified in other organs. Peri-epithelial fibroblasts, for example, resemble *COL15A1*$^{+}$ fibroblasts described in skin and gut.[57] APC-like fibroblasts share features with MHC-II^{+} populations observed in cancer,[60] synovium[61] and cardiac tissues.[62] Fibroblasts can express MHC under inflammation, but whether this directly regulates T cells remains context and tissue dependent.[8–10] Stress-responsive and inflammatory fibroblasts share transcriptional signatures with those in lung fibrosis,[63] inflammatory bowel disease,[64] and synovitis.[65] Our framework refines these into a niche-aware system applicable across health and disease, supporting comparisons in conditions where fibroblast plasticity and immune regulation are altered in disease.

Fibroblast-defined MCIMs were enriched for ligands mediating cytokine (*IL1B*, *TNF*), chemokine (*CXCL1/2/9/10/12*), and immune checkpoint (*PD-L1*, *ICAM1*) signaling. These modules mapped to immune-infiltrated foci or peri-epithelial zones, echoing the concept of parainflammation, a state where structural cells sustain low-level immune activity to preserve tissue homeostasis.[19,20] To spatially resolve these functional hubs, we developed a hist2omics pipeline linking spatial transcriptomics, multiplexed protein imaging, and same-slide histology. STARComm-inferred anatomically grounded spatial communication modules centered on para-inflammatory or immune-active fibroblasts. More granularly, this analysis revealed nominated fibroblast-centered therapeutic axes, including *IL1B*, *TNF*, *PD-L1*, and key chemokines, which highlight potential points of intervention across oral diseases and beyond. However, rather than pointing to a single drug readily translatable to the clinic, these findings should be viewed as a framework that nominates pathways of interest and generates hypotheses for future therapeutic exploration.

In our scRNA-seq and spatial-proteomics data, fibroblasts show antigen-presenting (MHC-I/II) and immunomodulatory roles, as previously described.[58–60] In cancer, fibroblast-driven immune exclusion and checkpoint signaling suppress anti-tumor immunity, creating immunologically cold microenvironments,[61–63] and, in chronic inflammation and autoimmunity,[64] modulating fibroblast-derived cues could limit tissue damage and aberrant remodeling without globally suppressing immune function.[61–63] Our atlas maps immune and fibroblast heterogeneity at single-cell and spatial resolution, enabling stromal reprogramming strategies for precision immunomodulation in cancer and chronic inflammation.[65,66] This pan-oral proteotranscriptomic atlas lays the groundwork for perturbation-guided, fibroblast-centric therapies that reprogram stromal-immune ecosystems to durably restore barrier integrity across the body.

Limitations of the study

Several important limitations highlight the need for continued investigation. Differences in cell-type composition inferred from the scRNA-seq analysis may be confounded by study-specific tissue processing, dissociation, and library-preparation protocols, which can systematically bias the recovery and relative abundance of certain cell types. Although dataset integration was used to harmonize transcriptional states across studies, such approaches do not fully disentangle protocol-driven effects from true biological variation in cellular proportions. Consequently, comparisons of cell-type composition were interpreted cautiously and treated as descriptive rather than quantitative or causal.

Our dataset is primarily derived from healthy adult tissues, limiting inference in pediatric, geriatric, and disease-specific contexts. Certain niches, such as peri-neural spaces and ectopic lymphoid sites, were underpowered for confident annotation. Transition zones and submucosa may harbor uncharacterized fibroblast states. MCIMs and predicted L-R interactions require functional validation, as spatial transcriptomics confirmed STARComm-inferred hubs but not L-R binding, co-localization, or downstream signaling.

Each platform has inherent limitations[67]: MERSCOPE, enables high-throughput transcript detection but is destructive to tissue architecture, and, while Xenium precisely assigning transcript position relative to cell boundaries, segmentation is still a challenge. These caveats highlight the need for orthogonal approaches, including spatial proteomics to capture protein-level signaling, histological overlays to preserve context, and *ex vivo* perturbation models (e.g., fibroblast subtype-specific targeting). Together, these complementary strategies will provide a more comprehensive and functionally grounded view of multicellular communication.

A central contribution of this study is the spatial inference of multicellular communication architectures rather than direct demonstration of physical L-R engagement. Using two independent spatial-transcriptomics platforms (MERSCOPE and Xenium), we observed convergent spatial patterning of fibroblast-centered ligand expression and immune cell proximity across oral mucosae and glands. While these orthogonal technologies differ substantially in chemistry, resolution, and tissue handling, both revealed recurrent, anatomically anchored communication hubs enriched for chemokine, cytokine, and checkpoint-associated signaling. Importantly, these observations reflect spatial co-enrichment and recurrence of signaling programs, not biochemical validation of L-R binding or downstream activation. STARComm therefore serves as a hypothesis-generating framework that identifies where and between which cell states communication is likely to occur, rather than asserting that signaling is functionally executed *in situ*. MCIMs were defined using pooled, multi-donor datasets to prioritize recurrent spatial programs conserved across patients, rather than cell-pair interactions driven by individual samples. The concordance of inferred modules across platforms and across donors supports the robustness of these spatial programs,

but functional validation will require perturbation-based approaches, including targeted fibroblast manipulation, receptor blockade, and *ex vivo* tissue models.

In summary, we present a comprehensive single-cell and spatial proteotranscriptomic atlas of adult human oral tissues, a harmonized fibroblast classification system, including two niche-partitioned stress-responsive fibroblast states, which represent spatially organized manifestations of conserved fibroblast stress programs rather than entirely novel cell types and influenced by their niche of origin, as well as modular tools to interrogate spatially resolved immunoregulation. By integrating histology, spatial multiomics, and computational inference, we lay the groundwork for targeted modulation of fibroblast activity in fibrosis, cancer, and autoimmunity. This foundational map enables future efforts to model, compare, and therapeutically manipulate structural-immune circuits across development, aging, and disease.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Kevin M. Byrd (byrdk6@vcu.edu).

Materials availability

This study did not generate new unique reagents. Antibodies, probe panels, and commercial kits used in this study are listed in the key resources table. Custom probe panels for spatial transcriptomics were designed in collaboration with commercial vendors and are described in the supplementary tables. Reasonable requests for additional information regarding materials will be addressed by the lead contact.

Data and code availability

- All data, including links to original raw data from each of the 11 studies (oral mucosa,[2,5,29–32] salivary glands,[31,33–35] and pulp[32,36,37]) can be found at GSE links: GSE217720; GSE164241; https://www.covid19cellatlas.org/; GSE161267; GSE152042; GSE201333; GSE180544; GSE199209; GSE188478; GSE146123; GSE185222. Original raw data for soft palate, labial mucosa, and parotid salivary gland can be also found at dbGaP: phs002446.v1.p1; GSE317219; GSE298850. The data can also be analyzed at https://cellxgene.cziscience.com/collections/065ad318-59fd-4f8c-b4b1-66caa7665409.
- Analysis notebooks and CELLxFEATURE matrices for MERSCOPE and PhenoCycler-Fusion 2.0 data are available at https://github.com/Locilab/Oral-Craniofacial-Atlas; https://doi.org/10.5281/zenodo.18474758.
- Any additional information required to reanalyze the data reported in this work is available from the lead contact upon request
- No restrictions apply to the reuse of the available codes.

ACKNOWLEDGMENTS

I.S. and K.M.B. want to extend a heartfelt thanks to the Human Cell Atlas Consortium, specifically Sarah Teichman, Aviv Regev, Ellen Todres, and Lucia Robson, as well as the HCA Oral & Craniofacial Bionetwork, for supporting this and ongoing projects to map the oral and craniofacial tissues in health and diseases. We also want to thank OMAPiX Inc, Psomagen, and Histoserv for their thoughtful support of this and future projects. Furthermore, we acknowledge that this article has become stronger and more comprehensive through conversations with Katarzyna M. Tyc (VCU) and with Di Wu and Ji-Eun Park, both at UNC. Figures 1D, S1C, and 4A were adapted using images from BioRender.com and released under a Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND) 4.0 International license. Figures 1A and S1B utilized and adapted the "salivary-glands" icon by Servier (https://smart.servier.com/), which is licensed under CC-BY 3.0 (https://creativecommons.org/licenses/by/3.0/). Data services in support of the research project were provided by the VCU Massey Comprehensive Cancer Center Bioinformatics Shared Resource. We would also like to thank the support and the staff of the NIDCR Office of the Clinical Director, NIDCR Dental Clinic, the National Eye Institute Ophthalmology Clinic, and the NCI Center for Cancer Research Anatomic Pathology. This work utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). We would like to thank the following NIH core facilities that enabled this study: NIDCR/NICHD Computational Biology and Genomics Core, the NIDCR Microscopy Core, the NIDCR Combined Technical Research Core, and the Laboratory of Genitourinary Cancer Pathogenesis (LCGP) Microscopy Core Facility. We would also like to thank the generous support of the NIH/UMD Graduate Partnership Program for TJFP. This research was principally supported through research awards to B.M.W. from the Division of Intramural Research (DIR) Program of the National Institute of Dental and Craniofacial Research of the National Institutes of Health (NIH/NIDCR ZIA: DE000704). Massey and this work have been supported in part with funding from an NIH-NCI Cancer Center Support Grant CA016059. D.P. is supported by Fundação para a Ciência e a Tecnologia (2020.08715.BD). I.S. is funded by Barts Charity (MGU045/G-001524), the Royal Society (RGS/R2/202291), and the British Skin Foundation (004/RA/23). This work was supported by generous start-up funds from the ADA Science & Research Institute (Volpe Research Scholar Award), Department of Oral and Molecular Craniofacial Biology, Philips Institute for Oral Health Research and VCU Massey Comprehensive Cancer Center (start-up funds), and NIH/NIDCR 1RM1DE035338 to K.M.B. P.R.T. was partly supported by the Chan Zuckerberg Initiative DAF award 2022-237918 and research grants from NIH/NHLBI (R01HL146557, R01HL160939, and R01HL153375). We want to thank the Core Unit for Systems Medicine at the University of Würzburg (supported by the IZKF Würzburg, project Z-6). This work was supported by funding from the German Cancer Aid (MSNZ Würzburg/NG3) and the European Union/European Research Council (ERC Starting Grant number 101042738/OralNiche) to K.K. This publication is part of the Human Cell Atlas: http://www.humancellatlas.org/publications.

AUTHOR CONTRIBUTIONS

B.F.M., K.L.A.H., D.P., I.S., J.L., and K.M.B. conceptualized the project. B.F.M., K.L.A.H., D.P., A.P., R.S., P.R.T., X.Z., M.K., S.A.T., J.L., and K.M.B. developed methods for data analysis. B.F.M., D.P., N.K., Q.T.E., B.T.R., A.G., S.H., L.S., M.B., P.P., Z.K., N.H., K.I.K., A.K., S.P., KK., B.M.W., I.S., and K.M.B. supported sample collection. B.F.M., K.L.A.H., D.P., X.Z., M.K., A.V.P., B.T.R., A.P., R.S., P.R.T., B.N., M.M.M., I.S., J.L., and K.M.B. performed experimental analysis. B.F.M. and K.M.B. wrote the original draft. B.F.M., K.L.A.H., D.P., A.P., X.Z., Q.T.E., S.A.T., R.S., P.R.T., I.S., J.L., and K.M.B. reviewed and edited the final manuscript.

DECLARATION OF INTERESTS

The authors had access to the study data and reviewed and approved the final manuscript. Although the authors view each of these as noncompeting financial interests, B.F.M., K.L.A.H., D.P., Q.T.E., A.V.P., K.I.K., S.A.T., A.K., S.P., K.K., B.W.M., I.S., J.L., and K.M.B. are all active members of the Human Cell Atlas; furthermore, S.A.T. has consulted for Roche and Genentech and is a scientific advisor for Biogen, GlaxoSmithKline, and Foresite Labs. I.S. is a consultant for L'Oréal Research and Innovation; K.M.B. is a scientific advisor at Arcato Laboratories; K.M.B. and J.L. are co-founders of Stratica Biosciences, Inc. K.M.B., J.L., K.L.A.H., and B.F.M. hold the patent for TACIT.

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
  - Ethics statement

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cpblue.2026.100007.

Received: December 1, 2024
Revised: September 23, 2025
Accepted: February 27, 2026
Published: March 23, 2026

REFERENCES

1. Byrd, K.M., and Gulati, A.S. (2021). The "Gum-Gut" Axis in Inflammatory Bowel Diseases: A Hypothesis-Driven Review of Associations and Advances. Front. Immunol. *12*, 620124. https://doi.org/10.3389/fimmu.2021.620124.
2. Squier, C.A., and Kremer, M.J. (2001). Biology of oral mucosa and esophagus. J. Natl. Cancer Inst. Monogr., *7*–15. https://doi.org/10.1093/oxfordjournals.jncimonographs.a003443.
3. Moutsopoulos, N.M., and Konkel, J.E. (2018). Tissue-specific immunity at the oral mucosal barrier. Trends Immunol. *39*, 276–287.
4. Peng, X., Cheng, L., You, Y., Tang, C., Ren, B., Li, Y., Xu, X., and Zhou, X. (2022). Oral microbiota in human systematic diseases. Int. J. Oral Sci. *14*, 14. https://doi.org/10.1038/s41368-022-00163-7.
5. Huang, N., Pérez, P., Kato, T., Mikami, Y., Okuda, K., Gilmore, R.C., Conde, C.D., Gasmi, B., Stein, S., Beach, M., et al. (2021). SARS-CoV-2 infection of the oral cavity and saliva. Nat. Med. *27*, 892–903.
6. Caetano, A.J., Human Cell Atlas Oral and Craniofacial Bionetwork; Sequeira, I., and Byrd, K.M. (2022). A Roadmap for the Human Oral and Craniofacial Cell Atlas. J. Dent. Res. *101*, 1274–1288. https://doi.org/10.1177/00220345221110768.
7. Presland, R.B., and Dale, B.A. (2000). Epithelial structural proteins of the skin and oral cavity: function in health and disease. Crit. Rev. Oral Biol. Med. *11*, 383–408. https://doi.org/10.1177/10454411000110040101.
8. Şenel, S. (2021). An Overview of Physical, Microbiological and Immune Barriers of Oral Mucosa. Int. J. Mol. Sci. *22*, 7821. https://doi.org/10.3390/ijms22157821.
9. Chibly, A.M., Aure, M.H., Patel, V.N., and Hoffman, M.P. (2022). Salivary gland function, development, and regeneration. Physiol. Rev. *102*, 1495–1552. https://doi.org/10.1152/physrev.00015.2021.
10. Gupta, S., Yamada, E., Nakamura, H., Perez, P., Pranzatelli, T.J., Dominick, K., Jang, S.I., Abed, M., Martin, D., Burbelo, P., et al. (2024). Inhibition of JAK-STAT pathway corrects salivary gland inflammation and interferon driven immune activation in Sjögren's disease. Ann. Rheum. Dis. *83*, 1034–1047. https://doi.org/10.1136/ard-2023-224842.
11. Regezi, J.A., Sciubba, J.J., and Jordan, R.C. (2016). Oral Pathology: Clinical Pathologic Correlations (Elsevier Health Sciences).
12. Moayedi, Y., Michlig, S., Park, M., Koch, A., and Lumpkin, E.A. (2021). Somatosensory innervation of healthy human oral tissues. J. Comp. Neurol. *529*, 3046–3061.
13. Byrd, K.M., Piehl, N.C., Patel, J.H., Huh, W.J., Sequeira, I., Lough, K.J., Wagner, B.L., Marangoni, P., Watt, F.M., Klein, O.D., et al. (2019). Heterogeneity within Stratified Epithelial Stem Cell Populations Maintains the Oral Mucosa in Response to Physiological Stress. Cell Stem Cell *25*, 814–829.e6. https://doi.org/10.1016/j.stem.2019.11.005.
14. Wu, R.-Q., Zhang, D.-F., Tu, E., Chen, Q.-M., and Chen, W. (2014). The mucosal immune system in the oral cavity—an orchestra of T cell diversity. Int. J. Oral Sci. *6*, 125–132.
15. Glim, J.E., Everts, V., Niessen, F.B., Ulrich, M.M., and Beelen, R.H.J. (2014). Extracellular matrix components of oral mucosa differ from skin and resemble that of foetal skin. Arch. Oral Biol. *59*, 1048–1055. https://doi.org/10.1016/j.archoralbio.2014.05.019.
16. Cavagnero, K.J., and Gallo, R.L. (2022). Essential immune functions of fibroblasts in innate host defense. Front. Immunol. *13*, 1058862.
17. Krausgruber, T., Fortelny, N., Fife-Gernedl, V., Senekowitsch, M., Schuster, L.C., Lercher, A., Nemc, A., Schmidl, C., Rendeiro, A.F., Bergthaler, A., and Bock, C. (2020). Structural cells are key regulators of organ-specific immune responses. Nature *583*, 296–302. https://doi.org/10.1038/s41586-020-2424-4.
18. Easter, Q.T., Fernandes Matuck, B., Beldorati Stark, G., Worth, C.L., Predeus, A.V., Fremin, B., Huynh, K., Ranganathan, V., Ren, Z., Pereira, D., et al. (2024). Single-cell and spatially resolved interactomics of tooth-associated keratinocytes in periodontitis. Nat. Commun. *15*, 5016. https://doi.org/10.1038/s41467-024-49037-y.
19. Easter, Q.T., Alvarado-Martinez, Z., Kunz, M., Matuck, B.F., Rupp, B.T., Weaver, T., Ren, Z., Tata, A., Caballero-Perez, J., Oscarson, N., et al. (2025). Polybacterial intracellular macromolecules shape single-cell inflammatory profiles in upper airway epithelia. NPJ Biofilms Microbiomes *11*, 100. https://doi.org/10.1038/s41522-025-00735-5.
20. Chen, M., and Xu, H. (2015). Parainflammation, chronic inflammation, and age-related macular degeneration. J. Leukoc. Biol. *98*, 713–725. https://doi.org/10.1189/jlb.3RI0615-239R.
21. Aran, D., Lasry, A., Zinger, A., Biton, M., Pikarsky, E., Hellman, A., Butte, A.J., and Ben-Neriah, Y. (2016). Widespread parainflammation in human cancer. Genome Biol. *17*, 145. https://doi.org/10.1186/s13059-016-0995-z.
22. Buechler, M.B., Pradhan, R.N., Krishnamurty, A.T., Cox, C., Calviello, A.K., Wang, A.W., Yang, Y.A., Tam, L., Caothien, R., Roose-Girma, M., et al. (2021). Cross-tissue organization of the fibroblast lineage. Nature *593*, 575–579. https://doi.org/10.1038/s41586-021-03549-5.
23. Gao, Y., Li, J., Cheng, W., Diao, T., Liu, H., Bo, Y., Liu, C., Zhou, W., Chen, M., Zhang, Y., et al. (2024). Cross-tissue human fibroblast atlas reveals myofibroblast subtypes with distinct roles in immune modulation. Cancer Cell *42*, 1764–1783.e10. https://doi.org/10.1016/j.ccell.2024.08.020.
24. Liu, K., Cui, Y., Han, H., Guo, E., Shi, X., Xiong, K., Zhang, N., Zhai, S., Sang, S., Liu, M., et al. (2025). Fibroblast atlas: Shared and specific cell types across tissues. Sci. Adv. *11*, eado0173. https://doi.org/10.1126/sciadv.ado0173.
25. Korsunsky, I., Wei, K., Pohin, M., Kim, E.Y., Barone, F., Major, T., Taylor, E., Ravindran, R., Kemble, S., Watts, G.F.M., et al. (2022). Cross-tissue, single-cell stromal atlas identifies shared pathological fibroblast phenotypes in four chronic inflammatory diseases. Med *3*, 481–518.e14. https://doi.org/10.1016/j.medj.2022.05.002.
26. Huynh, K.L.A., Tyc, K.M., Matuck, B.F., Easter, Q.T., Pratapa, A., Kumar, N.V., Pérez, P., Kulchar, R.J., Pranzatelli, T.J.F., de Souza, D., et al. (2025).

Deconvolution of cell types and states in spatial multiomics utilizing TACIT. Nat. Commun. *16*, 3747. https://doi.org/10.1038/s41467-025-58874-4.

27. Huynh, K.L.A., Matuck, B.F., de Souza, D., Zhang, X., Fatobene, G., Soares Junior, L.A.V., Ferraz da Silva, L.F., Rocha, V.G., Byrd, K.M., and Liu, J. (2025). STARComm Scalably Detects Emergent Modules of Spatial Cell-Cell Communication in Inflammation and Cancer. Preprint at bioRxiv. https://doi.org/10.1101/2025.08.07.669133.
28. Pranzatelli, T.J.F., Perez, P., Ku, A., Matuck, B., Huynh, K., Sakai, S., Abed, M., Jang, S.I., Yamada, E., Dominick, K., et al. (2024). GZMK+CD8+ T cells Target A Specific Acinar Cell Type in Sjögren's Disease. Preprint at Res Sq. https://doi.org/10.21203/rs.3.rs-3601404/v1.
29. Hermans, F., Bueds, C., Hemeryck, L., Lambrichts, I., Bronckaers, A., and Vankelecom, H. (2022). Establishment of inclusive single-cell transcriptome atlases from mouse and human tooth as powerful resource for dental research. Front. Cell Dev. Biol. *10*, 1021459. https://doi.org/10.3389/fcell.2022.1021459.
30. Williams, D.W., Greenwell-Wild, T., Brenchley, L., Dutzan, N., Overmiller, A., Sawaya, A.P., Webb, S., Martin, D., NIDCD/NIDCR Genomics and Computational Biology Core; and Hajishengallis, G., et al. (2021). Human oral mucosa cell atlas reveals a stromal-neutrophil axis regulating tissue immunity. Cell *184*, 4090–4104.e15. https://doi.org/10.1016/j.cell.2021.05.013.
31. Pagella, P., de Vargas Roditi, L., Stadlinger, B., Moor, A.E., and Mitsiadis, T.A. (2021). A single-cell atlas of human teeth. iScience *24*, 102405. https://doi.org/10.1016/j.isci.2021.102405.
32. Ko, K.I., DerGarabedian, B.P., Chen, Z., Debnath, R., Ko, A., Link, B.N., Korostoff, J.M., and Graves, D.T. (2023). Distinct fibroblast progenitor subpopulation expedites regenerative mucosal healing by immunomodulation. J. Exp. Med. *220*, e20221350. https://doi.org/10.1084/jem.20221350.
33. Caetano, A.J., Yianni, V., Volponi, A., Booth, V., D'Agostino, E.M., and Sharpe, P. (2021). Defining human mesenchymal and epithelial heterogeneity in response to oral inflammatory disease. eLife *10*, e62810. https://doi.org/10.7554/eLife.62810.
34. Tabula Sapiens Consortium*; Jones, R.C., Karkanias, J., Krasnow, M.A., Pisco, A.O., Quake, S.R., Salzman, J., Yosef, N., Bulthaup, B., Brown, P., et al. (2022). The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. Science *376*, eabl4896. https://doi.org/10.1126/science.abl4896.
35. Costa-da-Silva, A.C., Aure, M.H., Dodge, J., Martin, D., Dhamala, S., Cho, M., Rose, J.J., Bassim, C.W., Ambatipudi, K., Hakim, F.T., et al. (2022). Salivary ZG16B expression loss follows exocrine gland dysfunction related to oral chronic graft-versus-host disease. iScience *25*, 103592. https://doi.org/10.1016/j.isci.2021.103592.
36. Horeth, E., Bard, J., Che, M., Wrynn, T., Song, E.A.C., Marzullo, B., Burke, M.S., Popat, S., Loree, T., Zemer, J., et al. (2023). High-Resolution Transcriptomic Landscape of the Human Submandibular Gland. J. Dent. Res. *102*, 525–535. https://doi.org/10.1177/00220345221147908.
37. Chen, M., Lin, W., Gan, J., Lu, W., Wang, M., Wang, X., Yi, J., and Zhao, Z. (2022). Transcriptomic Mapping of Human Parotid Gland at Single-Cell Resolution. J. Dent. Res. *101*, 972–982. https://doi.org/10.1177/00220345221076069.
38. Krivanek, J., Soldatov, R.A., Kastriti, M.E., Chontorotzea, T., Herdina, A.N., Petersen, J., Szarowska, B., Landova, M., Matejova, V.K., Holla, L.I., et al. (2020). Dental cell type atlas reveals stem and differentiated cell types in mouse and human teeth. Nat. Commun. *11*, 4816. https://doi.org/10.1038/s41467-020-18512-7.
39. Opasawatchai, A., Nguantad, S., Sriwilai, B., Matangkasombut, P., Matangkasombut, O., Srisatjaluk, R., and Charoensawan, V. (2022). Single-cell transcriptomic profiling of human dental pulp in sound and carious teeth: a pilot study. Front. Dent. Med. *2*, 806294.
40. Osumi-Sutherland, D., Xu, C., Keays, M., Levine, A.P., Kharchenko, P.V., Regev, A., Lein, E., and Teichmann, S.A. (2021). Cell type ontologies of the Human Cell Atlas. Nat. Cell Biol. *23*, 1129–1135. https://doi.org/10.1038/s41556-021-00787-7.
41. Luecken, M.D., Büttner, M., Chaichoompu, K., Danese, A., Interlandi, M., Mueller, M.F., Strobl, D.C., Zappia, L., Dugas, M., Colomé-Tatché, M., and Theis, F.J. (2022). Benchmarking atlas-level data integration in single-cell genomics. Nat. Methods *19*, 41–50. https://doi.org/10.1038/s41592-021-01336-8.
42. Browaeys, R., Gilis, J., Sang-Aram, C., De Bleser, P., Hoste, L., Tavernier, S., Lambrechts, D., Seurinck, R., and Saeys, Y. (2023). MultiNicheNet: a flexible framework for differential cell-cell communication analysis from multi-sample multi-condition single-cell transcriptomics data. Preprint at bioRxiv. https://doi.org/10.1101/2023.06.13.544751.
43. Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat. Protoc. *15*, 1484–1506. https://doi.org/10.1038/s41596-020-0292-x.
44. Börner, K., Blood, P.D., Silverstein, J.C., Ruffalo, M., Satija, R., Teichmann, S.A., Pryhuber, G.J., Misra, R.S., Purkerson, J.M., Fan, J., et al. (2025). Human BioMolecular Atlas Program (HuBMAP): 3D Human Reference Atlas construction and usage. Nat. Methods *22*, 845–860. https://doi.org/10.1038/s41592-024-02563-5.
45. Stringer, C., and Pachitariu, M. (2025). Cellpose3: one-click image restoration for improved cellular segmentation. Nat. Methods *22*, 592–599. https://doi.org/10.1038/s41592-025-02595-5.
46. Steele, L., Olabi, B., Roberts, K., Mazin, P.V., Koplev, S., Tudor, C., Rumney, B., Admane, C., Jiang, T., Correa-Gallegos, D., et al. (2025). A single-cell and spatial genomics atlas of human skin fibroblasts reveals shared disease-related fibroblast subtypes across tissues. Nat. Immunol. *26*, 1807–1820. https://doi.org/10.1038/s41590-025-02267-8.
47. Liu, Z., Bian, X., Luo, L., Björklund, Å.K., Li, L., Zhang, L., Chen, Y., Guo, L., Gao, J., Cao, C., et al. (2025). Spatiotemporal single-cell roadmap of human skin wound healing. Cell Stem Cell *32*, 479–498.e8. https://doi.org/10.1016/j.stem.2024.11.013.
48. Almet, A.A., Yuan, H., Annusver, K., Ramos, R., Liu, Y., Wiedemann, J., Sorkin, D.H., Landén, N.X., Sonkoly, E., Haniffa, M., et al. (2023). A Roadmap for a Consensus Human Skin Cell Atlas and Single-Cell Data Standardization. J. Invest. Dermatol. *143*, 1667–1677. https://doi.org/10.1016/j.jid.2023.03.1679.
49. Sikkema, L., Ramírez-Suástegui, C., Strobl, D.C., Gillett, T.E., Zappia, L., Madissoon, E., Markov, N.S., Zaragosi, L.E., Ji, Y., Ansari, M., et al. (2023). An integrated cell atlas of the lung in health and disease. Nat. Med. *29*, 1563–1577. https://doi.org/10.1038/s41591-023-02327-2.
50. Wang, H., Yang, L., Chen, W., Li, K., Xu, M., Peng, X., Li, J., Zhao, F., and Wang, B. (2024). High-resolution subtyping of fibroblasts in gastric cancer reveals diversity among fibroblast subsets and an association between the MFAP5-fibroblast subset and immunotherapy. Front. Immunol. *15*, 1446613. https://doi.org/10.3389/fimmu.2024.1446613.
51. Liu, X., Dai, K., Zhang, X., Huang, G., Lynn, H., Rabata, A., Liang, J., Noble, P.W., and Jiang, D. (2023). Multiple Fibroblast Subtypes Contribute to Matrix Deposition in Pulmonary Fibrosis. Am. J. Respir. Cell Mol. Biol. *69*, 45–56. https://doi.org/10.1165/rcmb.2022-0292OC.
52. Pratapa, A., Mansouri, S., Nikulina, N., Matuck, B., Schneider, M.A., Byrd, K.M., Savai, R., Tata, P.R., and Singh, R. (2025). SAME: Topology-flexible transforms enable robust integration of multimodal spatial omics. Preprint at bioRxiv. https://doi.org/10.1101/2025.07.12.664419.
53. Gatenbee, C.D., Baker, A.M., Prabhakaran, S., Swinyard, O., Slebos, R.J.C., Mandal, G., Mulholland, E., Andor, N., Marusyk, A., Leedham, S., et al. (2023). Virtual alignment of pathology image series for multi-gigapixel whole slide images. Nat. Commun. *14*, 4502. https://doi.org/10.1038/s41467-023-40218-9.
54. Hu, Y., Rong, J., Xu, Y., Xie, R., Peng, J., Gao, L., and Tan, K. (2024). Unsupervised and supervised discovery of tissue cellular neighborhoods from cell phenotypes. Nat. Methods *21*, 267–278. https://doi.org/10.1038/s41592-023-02124-2.

55. Ding, D.Y., Tang, Z., Zhu, B., Ren, H., Shalek, A.K., Tibshirani, R., and Nolan, G.P. (2025). Quantitative characterization of tissue states using multiomics and ecological spatial analysis. Nat. Genet. *57*, 910–921. https://doi.org/10.1038/s41588-025-02119-z.

56. Taylor, P.C., Weinblatt, M.E., McInnes, I.B., Atsumi, T., Strand, V., Takeuchi, T., Bracher, M., Brooks, D., Davies, J., Goode, C., et al. (2023). Anti-GM-CSF otilimab versus sarilumab or placebo in patients with rheumatoid arthritis and inadequate response to targeted therapies: a phase III randomised trial (contRAst 3). Ann. Rheum. Dis. *82*, 1527–1537. https://doi.org/10.1136/ard-2023-224449.

57. Lendahl, U., Muhl, L., and Betsholtz, C. (2022). Identification, discrimination and heterogeneity of fibroblasts. Nat. Commun. *13*, 3409. https://doi.org/10.1038/s41467-022-30633-9.

58. Geppert, T.D., and Lipsky, P.E. (1985). Antigen presentation by interferon-gamma-treated endothelial cells and fibroblasts: differential ability to function as antigen-presenting cells despite comparable Ia expression. J. Immunol. *135*, 3750–3762.

59. Shenoy, A.T., Lyon De Ana, C., Arafa, E.I., Salwig, I., Barker, K.A., Korkmaz, F.T., Ramanujan, A., Etesami, N.S., Soucy, A.M., Martin, I.M.C., et al. (2021). Antigen presentation by lung epithelial cells directs CD4(+) T(RM) cell function and regulates barrier immunity. Nat. Commun. *12*, 5834. https://doi.org/10.1038/s41467-021-26045-w.

60. Fergatova, A., and Affara, N.I. (2023). The cellular triumvirate: fibroblasts entangled in the crosstalk between cancer cells and immune cells. Front. Immunol. *14*, 1337333. https://doi.org/10.3389/fimmu.2023.1337333.

61. Maru, S.Y., Wetzel, M., Mitchell, J.T., Gross, N.E., Andaloori, L., Howe, K., Kartalia, E., Mo, G., Leatherman, J., Ho, W.J., et al. (2025). Antigen-presenting cancer-associated fibroblasts in murine pancreatic tumors differentially control regulatory T cell phenotype and function via CXCL9 and CCL22. Preprint at bioRxiv. https://doi.org/10.1101/2025.03.27.645833.

62. Kerdidani, D., Aerakis, E., Verrou, K.M., Angelidis, I., Douka, K., Maniou, M.A., Stamoulis, P., Goudevenou, K., Prados, A., Tzaferis, C., et al. (2022). Lung tumor MHCII immunity depends on in situ antigen presentation by fibroblasts. J. Exp. Med. *219*, e20210815. https://doi.org/10.1084/jem.20210815.

63. Song, J., Wei, R., Liu, C., Zhao, Z., Liu, X., Wang, Y., Liu, F., and Liu, X. (2025). Antigen-presenting cancer associated fibroblasts enhance antitumor immunity and predict immunotherapy response. Nat. Commun. *16*, 2175. https://doi.org/10.1038/s41467-025-57465-7.

64. Lee, B., Lee, S.H., and Shin, K. (2022). Crosstalk between fibroblasts and T cells in immune networks. Front. Immunol. *13*, 1103823. https://doi.org/10.3389/fimmu.2022.1103823.

65. Lopes-Pacheco, M., and Rocco, P.R.M. (2023). Functional enhancement strategies to potentiate the therapeutic properties of mesenchymal stromal cells for respiratory diseases. Front. Pharmacol. *14*, 1067422. https://doi.org/10.3389/fphar.2023.1067422.

66. Akiyama, T., Yasuda, T., Uchihara, T., Yasuda-Yoshihara, N., Tan, B.J.Y., Yonemura, A., Semba, T., Yamasaki, J., Komohara, Y., Ohnishi, K., et al. (2023). Stromal Reprogramming through Dual PDGFRα/β Blockade Boosts the Efficacy of Anti-PD-1 Immunotherapy in Fibrotic Tumors. Cancer Res. *83*, 753–770. https://doi.org/10.1158/0008-5472.Can-22-1890.

67. Alexandrov, T., Saez-Rodriguez, J., and Saka, S.K. (2023). Enablers and challenges of spatial omics, a melting pot of technologies. Mol. Syst. Biol. *19*, e10571. https://doi.org/10.15252/msb.202110571.

68. Bankhead, P., Loughrey, M.B., Fernández, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., et al. (2017). QuPath: Open source software for digital pathology image analysis. Sci. Rep. *7*, 16878. https://doi.org/10.1038/s41598-017-17204-5.

69. Pringle, S., Wang, X., Verstappen, G.M.P.J., Terpstra, J.H., Zhang, C.K., He, A., Patel, V., Jones, R.E., Baird, D.M., Spijkervet, F.K.L., et al. (2019). Salivary Gland Stem Cells Age Prematurely in Primary Sjögren's Syndrome. Arthritis Rheumatol. *71*, 133–142. https://doi.org/10.1002/art.40659.

70. Kanemaru, K., Cranley, J., Muraro, D., Miranda, A.M.A., Ho, S.Y., Wilbrey-Clark, A., Patrick Pett, J., Polanski, K., Richardson, L., Litvinukova, M., et al. (2023). Spatially resolved multiomics of human cardiac niches. Nature *219*, 801–810. https://doi.org/10.1038/s41586-023-06311-1.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Human TruStain FcX	BioLegend	Cat# 422301
Multiplex IF antibody panel	Akoya Biosciences	See Table S6
TotalSeq-B0251	BioLegend	Custom
TotalSeq-B0252	BioLegend	Custom
Zombie Violet Fixable Viability Stain	Thermo Fisher	Cat# L34955
Biological samples		
FFPE oral mucosa and salivary gland blocks	UNC School of Medicine	UNC IRB 22-1786
Labial mucosa samples	NIH	IRB 15-D-0051
Parotid salivary gland biopsies	University of Groningen	METc 2008.078; METc 2023.281
Chemicals, peptides, and recombinant proteins		
Acridine orange/propidium iodide	Logos Biosystems	F23001
Advanced DMEM F12	Gibco	Cat# 12634-010
Collagenase D	Sigma-Aldrich	Cat# 11088858001
Collagenase P	Sigma-Aldrich	Cat# 11213857001
CryoStor CS10	StemCell Technologies	Cat# 07930
DNase I	Sigma-Aldrich	Cat# 10104159001
EDTA	Sigma-Aldrich	Cat# E6758
Fetal Bovine Serum (FBS)	Gibco	Cat# 26140-079
GlutaMAX	Gibco	Cat# 35050-061
HEPES	Gibco	Cat# 15630-080
HistoChoice MB Tissue Fixative	VWR International	Cat# 16050-122
Methanol (MeOH)	Sigma-Aldrich	Cat# 322415
Paraformaldehyde (PFA)	Sigma-Aldrich	Cat# P6148
Penicillin/Streptomycin	Gibco	Cat# 15140-122
ROCK inhibitor Y-27632	Tocris	Cat# 1254
Tissue-Tek O.C.T. Compound	Sakura	Cat# 4583
Trypsin	Sigma-Aldrich	Cat# T4799
Critical commercial assays		
Chromium Next GEM Single Cell 3′ Kit v3.1	10× Genomics	PN-1000128
Chromium i7 Sample Index Kit	10× Genomics	PN-120262
MERSCOPE Verification Kit	Vizgen	PN-91600004
Miltenyi Multi-tissue Dissociation Kit A	Miltenyi Biotec	Cat# 130-110-201
Deposited data		
Integrated oral craniofacial atlas	cellxgene.cziscience.com	https://cellxgene.cziscience.com/collections/065ad318-59fd-4f8c-b4b1-66caa7665409
MERFISH probe gene panel	Vizgen	Table S5
Published oral mucosa and salivary gland scRNA-seq datasets	GEO	https://www.ncbi.nlm.nih.gov/geo/
This study scRNA-seq (soft palate, labial mucosa, parotid gland)	GEO	see resource availability
Xenium probe gene panel	10× Genomics	Custom panel; Table S4
Experimental models: Organisms/strains		
Human patient-derived samples	This study	See IRB approvals

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
AstroSuite (TACIT, Constellation, STARComm, Astrograph)	This study	–
CellChat v2	Jin Lab	https://github.com/sqjin/CellChat
CellPhoneDB v5	Vento-Tormo Lab	https://github.com/ventolab/CellphoneDB
Cellpose 3	Stringer and Pachitariu[45]	https://github.com/MouseLand/cellpose
Drug2Cell	Saeys Lab	https://github.com/saeyslab/drug2cell
Harmony	Korsunsky et al.[25]	https://github.com/immunogenomics/harmony
MultiNicheNet	Saeys Lab[42]	https://github.com/saeyslab/multinichenetr
QuPath	Bankhead et al.[68]	https://qupath.github.io
STARsolo	Wellcome Sanger Institute	https://github.com/cellgeni/STARsolo
Seurat v5.1.0	Satija Lab	https://satijalab.org/seurat
Trailmaker	Biomage/Parse Biosciences	https://scp.biomage.net
scDblFinder	Bioconductor	https://bioconductor.org/packages/scDblFinder
scIB pipeline	Theis Lab	https://github.com/theislab/scib

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

This study analyzed human oral and craniofacial tissues obtained from multiple institutional sources under approved protocols. Samples included oral mucosa, labial mucosa, soft palate, parotid salivary glands, and additional oral niches as specified in the STAR Methods. All samples were obtained either from diagnostic surgical resections, research biopsies, or pathology archives. Only tissues without malignant involvement or significant architectural distortion were included for atlas construction unless otherwise stated. Soft palate oral mucosa samples were obtained from surgical resections at MSNZ Würzburg. Labial mucosa samples were collected at the National Institutes of Health under an approved Central IRB protocol. Parotid salivary gland biopsies were obtained at the University of Groningen from patients undergoing elective head and neck dissection for oral squamous cell carcinoma, with confirmation that parotid tissue was free of tumor involvement. Archived FFPE samples used for spatial biology experiments were obtained from the UNC School of Medicine pathology archive and reviewed independently by two trained oral pathologists prior to inclusion.

Sex, age, and clinical metadata were recorded when available from the originating institutions. No samples were excluded based on sex. Inclusion criteria were defined by tissue quality, availability of sufficient material for downstream analyses, and compliance with institutional ethical approvals. Exclusion criteria included poor tissue preservation, extensive necrosis, or inadequate cellularity.

Ethics statement

All human samples were collected in accordance with institutional review board approvals at the respective institutions. Approved protocols included NIH Central IRB Protocol 15-D-0051, UNC IRB 22–1786, and University of Groningen IRB approvals METc 2008.078 and METc 2023.281. All applicable regulations concerning informed consent and the use of human biological material were followed. For archived and de-identified samples, consent requirements were waived as permitted by local regulations.

METHOD DETAILS

Published scRNA-seq harmonization

Raw FASTQ files from published single-cell RNA sequencing projects were retrieved using scripts from this GitHub repository: https://github.com/cellgeni/reprocess_public_10x. First, metadata were collected from the GEO soft family file, and the ENA web API was used to gather data format details (SRR/ERR) and link samples to runs. Raw reads were downloaded as SRA archives, 10× BAM files, or gzipped paired-end FASTQ files, and then converted to FASTQ using fastq-dump (SRA tools v2.11.0) or bamtofastq (v1.3.2). The raw reads were processed using STARsolo for mapping and quantification, with wrapper scripts from https://github.com/cellgeni/STARsolo/ to identify kit versions and sample specifics. The human reference genome matched Cell Ranger 2022-A standards. For 10× samples, STARsolo settings were optimized to mimic Cell Ranger v6 outputs, including UMI/barcode processing and paired-end alignment. Cell filtering was performed with EmptyDrops (Cell Ranger v4+), producing both exon-only and full-length gene counts as well as RNA velocity matrices.

Newly generated scRNA-seq data

Soft Palate (MSNZ Würzburg). Resected soft palate oral mucosa samples were placed in ice-cold Advanced DMEM+++ medium (Advanced DMEM F12 (Gibco) supplemented with 10 mmol/L HEPES (Gibco), 100 U/ml penicillin/streptomycin (Gibco),

1 × GlutaMax (Gibco), and 10 μM of the Rho kinase (ROCK) inhibitor compound Y-2763) and transferred to the laboratory on ice as soon as possible. For histology, representative pieces of tissue were embedded in Tissue-Tek O.C.T. compound (Sakura) and placed directly on dry ice. The O.C.T.-embedded tissues were stored at −80°C for long-term storage. O.C.T. blocks were cut into 10 mm sections and stained with hematoxylin and eosin (H&E, Morphisto) after PFA fixation. For cryopreservation, the tissues were cut into smaller pieces and were incubated in CryoStor CS10 (Stem Cell Technologies) freezing medium for 10 min on ice. Then, the samples were transferred to a −80°C freezer. Frozen samples were stored in −150°C long-term. Frozen tissue samples were thawed at 37°C and washed in pre-warmed Advanced DMEM+++ medium. After rinsing, the tissue pieces were incubated in Advanced DMEM+++ for 10 min at room temperature. Samples were minced to a size of 0.5–1 mm3 and digested in Advanced DMEM+++ supplemented with 2% fetal bovine serum (Gibco), 2 mg/mL collagenase-P (Sigma-Aldrich), 1 mg/mL collagenase-D (Sigma-Aldrich), 0.5 mg/mL DNase-I (Sigma-Aldrich) and 10 μmol/mL Y-27632 (Tocris) for 30 min in a thermomixer (Eppendorf) at 37°C, 1000 rpm. During incubation, the digestion mixture was vigorously resuspended every 10 min. Samples were further digested into single cells in 0.2% trypsin in PBS supplemented with 10 μmol/mL Y-27632 (Tocris). The digestion was stopped with EDTA. Single cell suspension was filtered through a pre-wetted Pluristrainer with a pore size of 70 μm. Antibody staining and sorting: Cells were stained with Zombie Violet Fixable Viability Stain (1:1000 in PBS, Thermo Fisher) and blocked with Human TruStain FcX Fc Receptor Blocking Solution (1:10 in PBS, Biolegend). Then, cells were stained with TotalSeqTM-B0251 and B0252 (1:50, Biolegend), and alive cells were sorted with a BD FACS Aria III flow cytometer and collected into a clean tube. Sorted samples were pooled, counted with Trypan Blue, and loaded onto a 10× Chromium Controller (10× Genomics). Gene expression and cell surface protein library preparation was performed according to the manufacturer's instructions for the 10× Chromium Next GEM Single Cell Library kit v3.1 Dual Index (10× Genomics). The libraries were sequenced on a NextSeq 2000 Illumina sequencer.

Labial Mucosa (National Institutes of Health). Samples were collected from patients who provided informed consent under the NIH Central IRB Protocol 15-D-0051 (Principal Investigator: Warner). The sections were immediately placed in ice-cold RPMI, dissected into 1–2-mm pieces, and dissociated using a Miltenyi Multi-tissue Dissociation Kit A. It was then placed in 10% formaldehyde for a 16-24h fixation at 4°C during the fixation step and subsequently placed in a −80 fridge. To initiate the thawing process, the Thaw Enhancer (10× Genomics PN-2000482) was incubated at 65°C for 10 min, followed by vortexing and a brief centrifugation to ensure no precipitate was present. The reagent was kept warm, and the absence of precipitate was verified before use. Thawed Enhancer was not kept on ice to prevent precipitation, and once thawed, it was maintained at 42°C for up to 10 min. Next, 0.1 volume of pre-warmed Enhancer was added to the fixed sample in Quenching Buffer. For example, 100μL of Enhancer was added to 1,000μL of fixed sample in Quenching Buffer, followed by pipetting to mix. Alternatively, to conserve Enhancer volume, cells were centrifuged at 850 RCF, 500μL of Quenching Buffer was removed, and 50μL of Enhancer was added to the sample. 50% glycerol was introduced to achieve a final concentration of 10%. For instance, 275μL of 50% glycerol was added to 1,100μL of fixed sample in Quenching Buffer and Enhancer, followed by pipetting to mix. The samples could be stored at −80° for up to 6 months. In the post-storage processing phase of our study, all the samples were processed at the same time and, the steps outlined below were followed: Samples were thawed at room temperature until no ice was present. The samples were then centrifuged at 850rcf for 5 min at room temperature. The supernatant was carefully removed without disturbing the pellet. The cell pellet was resuspended in 1 mL of 0.5× PBS +0.02% BSA* (optionally supplemented with 0.2 U/μL RNase Inhibitor) or Quenching Buffer and kept on ice. *RNase-free BSA was used at this step. The dissociation process was conducted at 37°C in an OctoMACS tissue disruptor using heated sleeves. Single-cell suspensions underwent serial filtration through 70- and 30-μm filters and were rinsed with 1× Hanks' buffered salt solution. Cell counting and viability assessments were performed using a Trypan blue exclusion assay, with suspensions having greater than 35% viability chosen for subsequent sequencing. For single-cell capture, library preparation, and sequencing, approximately 10,000 cells were targeted and loaded onto a 10× Genomics Chromium Next GEM Chip B. Following cell capture, library preparation was performed using the 10× Chromium Next GEM Single Cell 3′ kit v3, and the libraries were sequenced on a NextSeq500 sequencer (Illumina). Cell processing was undertaken using the 10× Genomics Chromium Controller and the Chromium Single Cell 3′ GEM, Library & Gel Bead Kit v3 (PN-1000075), adhering to the guidelines provided in the manufacturer's user guide. Viability and concentration assessments were conducted by staining aliquots of sorted cells with acridine orange and propidium iodide, followed by analysis using a fluorescent Cell Counter. Each sample was loaded with approximately 3,200 cells onto the Chromium Chip B, aiming for a recovery of 2,000 cells per sample during library preparation. The creation of gel beads in emulsion (GEMs) involved encapsulating single cells, reverse transcription reagents, and gel beads coated with barcoded oligos within an oil droplet. Reverse transcription was executed using a C1000 thermal cycler (Bio-Rad), resulting in complementary DNA (cDNA) libraries tagged with a cell barcode and unique molecular index (UMI). GEMs were then broken, and the cDNA libraries purified using Dynabeads MyOne SILANE (Invitrogen) with subsequent 12 amplification cycles. Amplified libraries underwent purification with SPRIselect magnetic beads (Beckman Coulter) and quantification using an Agilent Bioanalyzer High Sensitivity DNA chip (Agilent Technologies). Following steps included fragmentation, end repair, A-tailing, and double-sided size selection with SPRIselect beads. Ligation of Illumina-compatible adapters onto size-selected cDNA fragments occurred. Adapter-ligated cDNA was purified using SPRIselect beads, and uniquely identifiable indexes were incorporated during 12 amplification cycles. The finalized sequencing libraries underwent purification with SPRIselect beads, visualization using the Bioanalyzer High Sensitivity DNA chip, quantification with the KAPA SYBR FAST Universal qPCR Kit for Illumina (Roche) and StepOnePlus Real-Time PCR System (Applied Biosystems), normalization to 4nM, and pooling. Sequencing took place on a NextSeq 500 machine (Illumina), with libraries denatured and diluted following the standard Illumina protocol. A 1% PhiX sequencing control (Illumina) was spiked in, and the libraries were loaded onto the flow cell at 1.8pM.

Parotid Salivary Glands (University of Groningen). Biopsy samples of parotid SG tissue were obtained from donors after institutional review board [IRB] approval, (METc 2008.078, METc 2023.281) who were treated for a squamous cell carcinoma of the oral cavity. In these patients, an elective head and neck dissection procedure was performed. During this procedure, the parotid SG is exposed and removed. This tissue does not contain malignant cells, as oral squamous cell carcinoma does not disseminate to the parotid SG. Salivary gland tissue was processed into a cell suspension according to our previously published method[69], and CS10 cryopreservation media until use. After thawing, separate cells were processed using 10× Genomics protocol version 3.1 chemistry (PN 2000121) accompanying the Version 3.1 kit, with a targeted cell yield of 8000 per lane. Samples were not multiplexed per 10× machine lane. Sample indexes from the Chromium i7 Sample Index Kit (PN-120262) were used for identification of pooled sampled on sequencing chip. Sequencing was performed using the Next Seq2500 (RapidRun) sequencer for a total of 127 cycles.

Published scRNA-seq integration, pre-process and visualization, and spatial single-cell annotation using Trailmaker

scRNA-seq. The single-cell RNA sequencing dataset was analyzed and visualized using Trailmaker, which was first accessible on the Biomage community platform (Biomage; https://biomage.net/), now developed and maintained by Parse Biosciences. Pre-filtered count matrices were loaded onto the platform for additional processing. Barcodes were filtered in four steps: first, barcodes with fewer than 500 UMIs and those indicating low quality or signs of apoptosis (over 15% mitochondrial content) were excluded. A robust linear regression model from the 'MASS' package was then used to predict expected gene counts per barcode, applying a tolerance based on sample size to exclude outliers outside this range. Doublets were identified and removed using the 'scDblFinder' R package, filtering out barcodes with a score over 0.5. Following filtering, between 300 and 8,000 barcodes per sample remained. Data was normalized with a log transformation, and the top 2,000 most variable genes were identified using variance-stabilizing transformation. PCA was applied, with the first 40 components (capturing 95.65% of the variance) then batch-corrected using the 'Harmony' package. To benchmark data integration, we applied the scIB pipeline to evaluate five methods: Harmony, Scanorama, scVI, BBKNN, and FastMNN. Integration performance was assessed across both batch correction and biological conservation. For batch correction, we evaluated overall performance using principal component regression (PCR), average silhouette width (ASW), and graph-based iLISI scores. For biological conservation, we used cell cycle variance and normalized mutual information (NMI) to assess the preservation of biological structure. Louvain clustering and UMAP were conducted in Seurat for dimensionality reduction and cluster identification. Differentially expressed genes were identified for each cluster using the Wilcoxon rank-sum test ('presto' package). These samples were then imported into Trailmaker, bypassing initial filtering since pre-processing had been completed. The same analytical pipeline was applied to this subset as to the complete dataset. Cells were annotated based on literature and further classified with the CellTypist tool. For MERFISH data from MERSCOPE, segmented data, including barcodes, features, and matrices, was processed within Trailmaker. Cells with no transcripts were removed based on transcript count per cell, and cell types were annotated using the lasso tool, focusing on the most highly expressed genes within each Louvain cluster.

L-R analyses via Cellphone DB, CellChat, and MultiNicheNet

We employed CellPhoneDB[66] (v5, https://github.com/ventolab/CellphoneDB), leveraging its curated database of L-R interactions to analyze cell-to-cell signaling interactions at a single-cell level. Using default parameters, we calculated the likelihood of interactions between pairs of cell types based on the expression of known L-R pairs, resulting in probabilistic scores that highlight statistically significant intercellular communication. Additionally, we applied CellChat[67] (v2) to quantify signaling probabilities between manually defined sender and receiver cell groups. Both methods were applied to the same datasets, providing a comprehensive communication profile across different niches and enabling cross-validation of key interactions. We also employed MultiNicheNet,[42] a novel framework available at https://github.com/saeyslab/multinichenetr, to enhance the analysis of cell-cell communication within multi-sample multi-condition single-cell transcriptomics data, with a focus on oral mucosal tissues. The primary objectives of MultiNicheNet were to infer differentially expressed and active L-R pairs between conditions of interest and predict the putative downstream target genes of these pairs. The pipeline was applied in two different approaches, one related to a Tier 1 annotation dataset into the integrated oral craniofacial atlas. Statistical difference was employed to determine the top 50 interactions between receptor and ligands in glands versus mucosa. A second approach was made to determine the difference within each of the major niches, such as Salivary Glands and all Mucosa. Cells annotations were carried from Tier 2 to run downstream analysis. In the purpose of setting the threshold of cell, only cell types with more than 10 cells in all niches were included in the analysis. *p* values were extracted using muscat framework for DEG and testing which is the top highest expressed based on mixed linear models.

Sectioning for spatial biology

Samples from the pathology archive of UNC medicine school (UNC IRB 22–1786) were selected with two trained oral pathologists, only samples not containing architectural or cytologic alterations were included. FFPE Blocks were sectioned in 5 μm sequential sections and mounted in SuperFrost Plus (Thermo Fisher) slides for Phenocycler-Fusion 2.0, H&E and Stains. Coverslips (MERSCOPE slides) from MERSCOPE (Vizgen) were also used to mounting the sequential sections from each slide that went through other Phenocycler experiment using the '91600112 MERSCOPE User Guide Formalin-Fixed Paraffin-Embedded Tissue Sample Preparation RevB'. MERSCOPE TM Slides have a defined imageable of a 2 × 1.5cm rectangular where two sections from the same tissue were placed at time. All slides preparation was performed using RNA free water and RNAse protocol from Leica autotec system.

Multiplex protein immunofluorescence (formerly, CODEX; now, PhenoCycler fusion 2.0, Akoya Biosciences)

For multiplexed immunofluorescence (Multi-IF), we used the PhenoCycler Fusion 2.0 from Akoya Biosciences. All samples underwent deparaffinization using a gradual reduction of alcohol concentrations from 100% to 30%. Antigen retrieval was performed using AR9 (EDTA) buffer from Akoya Biosciences in a low-pressure cooker for 15 min. After cooling for 1 h, samples were hydrated with ethanol for 2 min followed by a 20-min incubation in staining buffer. The antibody cocktail buffer was prepared according to the Phenocycler Fusion Manual using four blockers and nuclease-free water in addition to the staining buffer. Antibodies (Table S6) were diluted to a 1:200 concentration in the antibody cocktail buffer. A mixed solution containing all primary antibodies was then prepared. Slides were incubated overnight at 4°C in a humidity chamber (Sigma-Aldrich) with the primary antibodies. After primary incubation the slides were placed in staining buffer for 2 min, followed by post-stain fixing solution (10% PFA in staining buffer) for 10 min. Following 3 baths of 2 min washing in 1× PBS, the slide was immersed in ice-cold MeOH for 5 min. The final fixative solution (FFS) was prepared in accordance with the Phenocycler manual. Slides were then incubated with the FFS for 20 min at room temperature. Following this, the slides underwent a PBS wash and were promptly transferred to the FCAD Machine (Akoya Biosciences) for flow cell mounting. Flow cell stitching was achieved through a 30-s application of high pressure to the top of the slides. Once mounted, the flow cells with the slides and sectioning were immersed in PhenoCycler (PCF) buffer for 10 min before being transferred to the Phenocycler Fusion. Reporters for each antibody were prepared using a Report Stock solution along with 5490 μL of nuclease-free water, 675 μL of 10× PCF buffer, 450 μL of PCF assay reagent, and 27 μL of in-house prepared concentrated DAPI to achieve a final DAPI concentration of 1:1000 per cycle. This process was repeated for a total of two slides per time. Reporters were included into a 1:50 dilution for each cycle, including the respective channels. 250 mL of each specific report stock solution were pipetted into a 96 well-plate that was sealed using an aluminum foil provided by Akoya Biosciences. Two slides were included into the fluid equipment at time, using the Phenocycler Fusion 2.0. Manual mapping was used to the scanning area to be scanned by the PhenoImager using brightfield. Low DMSO and High DMSO was prepared as suggested by the Index B from the Phenocycler Fusion 2.0 manual of operations.

Multiplexed error-robust fluorescence *in situ* hybridization (Vizgen MERSCOPE)

For multiplexed error-robust fluorescence *in situ* hybridization (MERFISH), we used MERSCOPE from Vizgen. To ensure maximum sensitivity and consistency, MERSCOPE verification was performed on representative samples from all oral niches using the MERSCOPE Verification Kit (Vizgen) according to the manufacturer's instructions (91600004 Rev D). Briefly, verification mimics all sample processing steps but uses a single-round smFISH readout of the EEF2 housekeeping gene to determine signal brightness, background fluorescence and consistent adhesion. Verification results with increased background in several tissue types made alterations to the manufacturer's protocol necessary. A harsh clearing protocol was developed and verified on additional samples in alignment with the Vizgen team. It replaced the usual digestion followed by single stage clearing of clearing resistant tissues with a multi-step protocol with increased Proteinase K (ProtK, NEB, Cat P8107S) concentrations. After gel embedding, samples were initially cleared with 1.25 mL clearing premix (Vizgen) and 120 μL ProtK for 4 h at 47°C. After incubation another 3.75 mL of clearing premix without ProtK was added to the samples and incubated over night at 47°C. Samples were washed three times for 5 min in Sample Prep Wash Buffer (Vizgen) to remove the clearing solution and digested in Digestion Mix (Vizgen) for 2 h at 37°C. 5 mL fresh Clearing Premix with 50 μL ProtK were added directly to the digestion reaction and incubated for 3 days at 37°C before continuing with autofluorescence quenching according to the unmodified MERSCOPE protocol. These modifications to the FFPE MERSCOPE protocol (91600112 Rev B) were used to generate all MERSCOPE results otherwise following the manufacturer's instructions. Samples were hybridized with MERSCOPE probes for the RNAs listed in Table S5 and smFISH probes for KRT14 (ENSG00000186847), KRT1 (ENSG00000167768), KRT6B (ENSG00000185479), APOD (ENSG00000189058), LORICRIN (ENSG00000203782) and B2M (ENSG00000166710) that were all part of the custom designed panel (Vizgen panel ID: BP0983). Next to transcript locations, tiff images were generated for DAPI and polyT as well as three proprietary cell boundary stains (Vizgen, PN 10400009) at 0.108 μm pixel size and the smFISH probes listed above.

Xenium spatial-transcriptomics and Histo2mics

For spatial transcriptomic profiling, FFPE tissue sections were processed using the Xenium *in situ* platform (10× Genomics) according to the manufacturer's user manual. A custom-designed probe panel targeting 300 genes of interest was used for transcript detection. Tissue sections were mounted onto Xenium-compatible slides and underwent the full automated workflow, including tissue permeabilization, probe hybridization, and signal amplification, as outlined in the protocol. Following the completion of the Xenium run, we opted not to proceed with the standard hematoxylin and eosin (H&E) staining. Instead, we performed the optional dequenching step described in the Xenium protocol, which enables removal of residual fluorescent signal and preserves the tissue for subsequent downstream applications, such as multiplexed spatial-proteomics. Following the dequenching step, the slides were placed in a storage solution consisting of 10% glycerol and 90% phosphate-buffered saline (PBS) to preserve tissue integrity and morphology. Slides were stored at 4°C in this buffer until further processing for spatial-proteomics, ensuring compatibility with downstream multiplexed immunofluorescence workflows. Slides were then subjected to a new round of antigen retrieval using EDTA-based buffer pH 9 (Akoya Biosciences) to prepare the tissue for multiplexed immunofluorescence. The same Phenocycler panel used in subsequent experiments was applied to ensure consistency across samples. After completion of the spatial-proteomics workflow, slides were immersed in HistoChoice MB Tissue Fixative (VWR International, Radnor, PA) for 24 h to facilitate flow cell removal. The flow cell was carefully detached using a #15 surgical blade, taking care not to damage the underlying tissue. Finally, slides were stained with hematoxylin and eosin (H&E) to assist in tissue segmentation and provide morphological context for downstream

analyses. The final output consisted of three spatially aligned images obtained from the same slide: Phenocycler multiplexed proteomics, Xenium spatial-transcriptomics, and H&E histology.

Cell Segmentation for Multi-IF and spatial-transcriptomics images using Cellpose3[47]

Cell segmentation was carried out using Cellpose3. We employed a pipeline that integrated both denoising and segmentation steps to improve the quality of input images and the accuracy of the segmentation outputs based in nuclei expansion. The model was custom-trained on H&E-stained tissue sections from the niches of the oral cavity, allowing for optimized performance on this tissue type, which often presents significant histological heterogeneity. Prior to segmentation, a denoising step was applied to reduce image noise and artifacts. The cell expansion was based on the auto-calibration tool from Cellpose3 and the starting trained model used was the Cyto2. For this analysis, we used the QuPath extension for Cellpose, allowing us to streamline the workflow by combining QuPath's visualization tools with the Cellpose segmentation model trained. Annotations created in QuPath were exported as a CSV file. The segmentation results were then used as input for further downstream analyses. Specifically, the segmented data were inputted into PCF and MERSCOPE platforms for spatial-transcriptomics and spatial-proteomics analysis, enabling the identification of spatial patterns and cell-to-cell interactions at single-cell resolution. We Inputted the segmentation mode in the Vizualizer tool (Vizgen), which provided an intuitive interface for examining the quality and accuracy of the segmentations. This visualization step was critical for ensuring that the segmentations aligned with biological expectations and tissue architecture.

AstroSuite for auto assignment of cell identities and states, TCNs, and MCIMs

AstroSuite contains interconnected algorithms for analyzing spatial biology; two of the algorithms, TACIT and Astrograph, have been published previously.[26] TACIT (Threshold-based Assignment of Cell Types from Multiplexed Imaging Data) for cell type annotation. We used TACIT for cell type annotation in our spatial omics dataset. TACIT is an unsupervised algorithm designed for single-cell resolved spatial modalities, including spatial-transcriptomics and proteomics. TACIT takes as input the normalized CELLxFEATURE matrix as a result of Cellpose3 cell segmentation and quantification of individual cell features such as probe intensity or count values. Additionally, a TYPExMARKER matrix, informed by expert knowledge, is used to indicate the relevance of specific markers for defining cell types. The annotation process occurs in two stages. In the first stage, cells are grouped into Microclusters (MCs) using the Louvain algorithm to capture highly homogeneous cell populations. Simultaneously, Cell Type Relevance scores (CTRs) are calculated for each cell through a linear combination of its normalized marker intensity vector with predefined cell type signatures. A higher CTR indicates a stronger association with a specific cell type and vice versa. Next, TACIT determines a threshold that distinguishes true positive cell type signals from background noise. CTRs are ranked, and segmental regression is applied to divide the CTR growth curve into 2 to 4 segments, identifying high and low relevance clusters. Unavoidably, there exists cells with low relevance score present in high relevance group and vice versa. An absolute CTR threshold is then determined to minimize such inconsistent assignments. Subsequently, cells with CTRs exceeding this threshold are labeled as positive for the associated cell type. In cases where cells are labeled with multiple cell types, TACIT employs a deconvolution step using the k-nearest neighbors (k-NN) algorithm and cells with unique cell type assignment to resolve such ambiguity.

Constellation for Tissue Cellular Neighborhood Assignment in Whole Slide Imaging. We developed Constellation to address the challenges of unsupervised Tissue Cellular Neighborhood (TCN) assignment in Whole Slide Imaging (WSI) data, building upon the original CytoCommunity framework. While CytoCommunity effectively clusters cells within smaller tissue slides, it struggles with scalability for large WSIs or multiple slides. Constellation introduces novel partitioning strategies and a consensus-based assignment approach, enabling efficient processing of large datasets. The first step is WSI partitioning, where tissues are divided into smaller, non-overlapping sub-regions to reduce computational complexity. The tissue, represented as a set of cells W, with spatial coordinates and phenotypic labels, is partitioned using three strategies: vertical stripes (based on x-coordinate boundaries), horizontal stripes (based on y-coordinate boundaries), and square grids (dividing along both axes). Each partition contains a minimum and maximum number of cells to ensure computational efficiency and biological relevance, where partitions exceeding the maximum cell count are subdivided, and smaller ones are merged. Formally, let $W = \bigcup_{i=1}^{K} S_i$, where S_i representing the *i-th* subset of cells, and K is the total number of partitions such that $S_i \cap S_j = \varnothing$ for all $i \neq j$. Next, sub-graph $G_i=(V_i,E_i)$ are constructed for each partition S_i capturing both phenotypic and spatial relationships among the cells. Here, V_i denotes the set of nodes (cells) withinS_i, where each node $v_n \in V_i$ is represented as a one-hot encoded phenotype vector:

$$v_n = [t_{n1}, t_{n2}, \ldots, t_{nT}] \text{ with } t_{ni} = 1 \text{ if phenotype } = i$$

and T is the total number of phenotypes. The edge set E_i connects nodes based on the k-nearest neighbors determined by Euclidean distance in the spatial coordinate space. The sub-graph is represented by an adjacency matrix $A_{G_i} \in \mathbb{R}^{|V_i| \times |V_i|}$, where entries are binary (1 for an edge, 0 otherwise), and a feature matrix $F_{G_i} \in \mathbb{R}^{|V_i| \times T}$, representing cell phenotypes. The collective unsupervised training stage processes these sub-graphs in a graph neural network (GNN). Each sub-graph is treated as an independent training sample, and the model processes them in mini-batches with a gradient accumulation technique over $N_{acc} = 32$ batches. The GNN minimizes a combined loss function for training:

$$L_t = \frac{1}{N_{acc}} (L_{MinCut} + L_{orth}),$$

where L_{MinCut} encourages separation between different clusters andL_{orth} promotes orthogonality, ensuring that each cell is confidently assigned to distinct TCNs. Finally, TCN assignment is performed using a consensus-based approach that aggregates results across the partitioning strategies. For each cell v_n, TCN assignments from vertical stripes, horizontal stripes, and square grid partitioning methods are compared, with the final assignment determined by majority voting:

$$c^n_{final} = mode\left(c^n_1, c^n_2, c^n_3\right),$$

where c^n_m denotes the TCN assignment from the m-th partitioning method. In rare cases where all three assignments differ, the cell is excluded to maintain result reliability. This consensus-based approach ensures robust TCN assignments by integrating diverse spatial perspectives within the WSI. STARComm for spatial-resolved cell-cell communication: In STARComm (*spatial r*eceptor-ligand analysis for cell-cell *comm*unication) analyses, we focus on identifying and characterizing co-location communication between cells based on 50 L-R interactions in spatial-transcriptomics datasets, such as MERSCOPE. The process begins by identifying cell pairs where communication may occur. Specifically, a cell that expresses a ligand gene (with value above 0 count) interacts with a nearby cell that expresses the corresponding receptor gene (also above 0 count), provided they are within 50 μm of each other in the spatial tissue layout. This process is repeated for all L-R pairs (50 pairs in total), resulting in a network of interactions where each pair of cells that meets the L-R criteria is connected by an edge. Once all the L-R interactions are established, we organize the spatial region into a grid of bins size 100 μm x 100 μm. Each bin in this grid captures the spatial region, and we ensure that each bin contains a very small number of edges representing L-R pairs. For each bin in the grid, we tally the count of individual L-R pairs that fall within that spatial region, providing a cumulative measure of communication strength in that area. To analyze the distribution of these interactions more deeply, we calculate the kernel density for each L-R pair.

$$\widehat{f}(x,y) = \frac{1}{n * h_x * h_y} \sum_{(x_i,y_i) \in n} K\left(\frac{x - x_i}{h_x}\right) K\left(\frac{y - y_i}{h_y}\right)$$

where $\widehat{f}(x,y)$ is the density at the grid (x, y); n is the total number of grids; h_x and h_y are the bandwidths in the x and y directions (default 10); (x_i,y_i) are the coordinates of the grid; and $K(.)$ is the kernel function - $K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$.

The kernel density estimation reveals how communication is spatially distributed—areas where the interactions between cells are more frequent will have higher density values, indicating that those locations are hubs of L-R communication. Finally, we apply the Louvain clustering algorithm to the grid, clustering the bins based on their communication density vectors. This clustering process groups spatial regions that exhibit similar co-location of L-R pairs meaning that the L-R interactions are concentrated in specific areas of the tissue. By clustering the bins, we capture and visualize how cell communication is spatially organized, highlighting areas where L-R interactions are occurring at higher densities, providing insight into the spatial dynamics of cellular communication across the tissue.

Astrograph for spatial data visualization. We developed Astrograph, an R Shiny application that enables interactive visualization and analysis of spatial cell type data. The app allows users to customize cell type annotations with user-defined colors and select specific tissues or regions for focused visualization. It supports the generation of heatmaps that display cell type distributions based on marker expression, which can be tailored to reflect either cell type-specific markers or cell state markers. Beyond visualizing cell types, Astrograph also facilitates the exploration of cell states, providing insights into the proportions of various cell types and states across tissues or regions. Additionally, the app offers Voronoi plot functionality to depict spatial relationships between cells, thereby visualizing cell neighborhoods and interactions. Furthermore, Astrograph supports quantitative analysis of cell type interactions by counting neighboring relationships and identifying patterns in spatial cell dynamics, offering a comprehensive approach to spatial omics data exploration.

Adapting Drug2Cell for spatial datasets

In the Spatial Drug2Cell[70] analysis, we apply the Drug2Cell approach to the MERSCOPE data after the cells have been annotated with their respective cell types. Once cell type annotation is completed, we proceed by extracting the drug scores for each cell after we run drug2cell packages. These drug scores quantify the response or relevance of individual cells to specific drugs, based on expression profiles and other molecular characteristics. After obtaining the drug scores, we map these scores onto the spatial coordinates of the cells, allowing us to visualize the distribution of drug responses across the tissue landscape. This spatial mapping enables us to analyze how drug sensitivity or resistance varies based on the physical location of the cells within the tissue. Furthermore, we can compare the drug responses between different cell types and regions, highlighting variations in drug effectiveness or targeting based on cellular identity and spatial context. This comparison can reveal how certain cell types, such as immune cells, fibroblasts, or tumor cells, may exhibit distinct drug responses depending on their location within the tissue microenvironment. The combination of spatial information with drug scores provides a comprehensive view of how drugs interact with different cell populations, aiding in targeted therapeutic strategies.

Motif graphs

The motif graph is a network-based representation designed to capture spatial adjacency relationships between groups of functional neighborhoods (FNs). A functional neighborhood can represent a specific TCN from Mega-CytoCommunity or an MCIM from

StarComm. In this graph, each node corresponds to an FN, and an edge exists between two nodes if there exist neighboring cells linking two different FNs, i.e., a cell from one FN is within 50 micrometers of a cell in the other FN. Since FNs may appear in multiple locations on a slide as instances, the motif graph effectively summarizes the interaction relationships between these types of FNs, with repeated FN motifs forming a subgraph structure. For each pair of connected FNs we compute the proportion of neighboring cells linking them relative to the total number of cells in both FNs. This metric reflects the frequency of neighboring relationships, and it is visually represented by the thickness of the edges in the motif graph. Note that in the visualization of motif graph, we highlight only the top three most frequent interactions between clusters to emphasize key connections.

GO and MSigDB pathway analysis

For all single-cell RNA-seq analyses, differentially expressed genes (DEGs) were identified using Trailmaker with a fold-change threshold of $\geq$0.5 and adjusted p value <0.05 unless otherwise noted. DEGs were filtered for each fibroblast subtype or disease state comparison and subjected to Gene Ontology (GO) and MSigDB pathway enrichment using the Enrichr platform (https://maayanlab.cloud/Enrichr/). Only significantly enriched terms (adjusted $p < 0.05$) were retained for visualization and interpretation. For Figure 4 (niche-specific fibroblast programs) and Figure 6 (chronic periodontitis), analyses focused on identifying subtype-specific pathways related to ECM organization, immune surveillance, cytokine signaling, and stress response. For Figure 6, pathway enrichment was performed using a more stringent DEG threshold (fold-change $\geq$1.0, adjusted p value <0.05) to reflect high-confidence spatial proteogenomic differences. Interactivenn (http://www.interactivenn.net/) was used to assess gene list overlap and visualize shared and unique gene sets between fibroblast subtypes across health and disease conditions.

QUANTIFICATION AND STATISTICAL ANALYSIS

General methods

A variety of tools and statistical approaches were used for data analysis, as appropriate for the data type and study objectives. Non-sequencing-based data were primarily analyzed using QuPath and/or Prism 9. The choice of statistical tests is detailed in the main text and figure legends. All statistical analyses were two-sided, and significance was determined at a threshold of (p value < 0.05). For comparisons between groups, a t test was applied when normality assumptions were met. If the data did not meet these assumptions, a Wilcoxon rank-sum test (Mann-Whitney U test) was used instead. All visualizations, including graphs and supplementary figures, were created using Prism 9/10, a community instance of Trailmaker (hosted by Parse), and the 'ggplot2', 'pheatmap', 'ggforce' package in R unless explicitly stated otherwise. Venn Diagrams were generated using http://www.interactivenn.net. In spatial proteotranscriptomics analyses, workflows for cell segmentation, cell type annotation, and tissue structure annotation were tailored to the data and the biological questions of interest. Cell segmentation was performed using the advanced capabilities of Cellpose3, which excels at robust segmentation of complex tissue images. For annotation, AstroSuite was employed, providing an automated pipeline for assigning cell identities, functional states, tissue structure neighborhoods, and multiplexed cell interaction maps (MCIMs). These methods are described in detail in the ''Cell Segmentation for Multi-IF and MERFISH Images using Cellpose3 and AstroSuite for Auto Assignment of Cell Identities and States, TCNs, and MCIMs'' sections. Heatmaps for spatial data were generated by summarizing marker expression. For PhenoCycler datasets, the mean expression levels were used to capture trends across the data, while for MERFISH datasets, the median count values were used to better handle the sparsity inherent in these datasets. When alternative methods for summarization were used, these are specified in the corresponding figure legends to ensure transparency and reproducibility. Differential expression analyses for genes, markers, and drug targets were conducted using the Seurat pipeline (version 5.1.0). This pipeline integrates preprocessing, normalization, and differential expression analysis, allowing us to compare expression profiles across groups effectively.