Single-cell RNA sequencing identifies heterogeneous cell subtypes within gingival tissue

Shujiao Qian
Shanghai Ninth People's Hospital
Qianru Huang
Shanghai Institute of Immunology
Junyu Shi
Shanghai Ninth People's Hospital
Linyi Zhou
Shanghai Ninth People's Hospital
Yi Zhao
Shanghai Institute of Immunology
Bin Li
Shanghai Institute of Immunology https://orcid.org/0000-0002-7640-8884
Hongchang Lai (hongchanglai@126.com)
Shanghai Ninth People's Hospital

Article

Keywords: single-cell sequencing, gingival, periodontitis

Posted Date: November 6th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-101475/v1

Single-cell RNA sequencing identifies heterogeneous cell subtypes within gingival tissue

Authors:

Shu-jiao Qian[1*], Qian-ru Huang[2,3*], Jun-yu Shi[1], Lin-yi Zhou[1], Yi Zhao[2,3], Bin Li[2,3#], Hong-chang Lai[1#]

[*]Authors contributing equally as co-first authors.

[#] Correspondence authors. Hong-chang Lai, lhc9@hotmail.com and Bin Li, binli@shsmu.edul.cn

[1]Department of Implant Dentistry, Shanghai Ninth Peoples' Hospital, College of Stomatology, Shanghai Jiao Tong University School of Medicine, National Clinical Research Center for Oral Diseases, Shanghai Key Laboratory of Stomatology & Shanghai Research Institute of Stomatology, Shanghai, 200011, China

[2]Shanghai Institute of Immunology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

[3]Department of Immunology and Microbiology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

Conflict of interest and source of funding statements

None of the authors have conflict of interest. This study has been supported by The Project of Biobank from Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine (YBKA201906).

Abstract

Owing to the high prevalence of periodontitis and its impact on quality of life, a rigorous characterization of periodontal tissue is necessary. However, this has previously been impeded by the cellular heterogeneity of periodontal tissue. Here, we characterized the gene expression of gingival tissue in health and periodontitis at single-cell resolution. We reported the loss of epithelial homeostasis-associated genes expressed by specific epithelial cell types in periodontitis. We detected endothelial cells highly associated with immune regulation and fibroblasts potentially involved in bone remodeling. We revealed the myeloid cells with diverse clusters dominated by pro-inflammatory phenotype. We further identified cell subtypes specifically expressing periodontitis related genes and highlighted the cell-cell communication alteration between conditions. This novel single-cell study generated a comprehensive catalog of cells types and interaction networks in the human gingiva for the first time, thus offering deeper insights into the biological foundation of periodontal homeostasis and diseases, which will be helpful in advancing periodontal diagnosis and therapy.

Keywords single-cell sequencing, gingival, periodontitis

Introduction

Periodontium is an umbrella term for tissues that surround and support the teeth, including soft tissues (gingiva and periodontal ligament) and hard tissues (alveolar bone and cementum)[1,2]. Periodontitis, a chronic or aggressive inflammatory lesion of the collective periodontium, is characterized by irreversible and progressive degradation of the periodontal tissue and causes tooth loss and alveolar bone defects. Current evidence has demonstrated independent associations between periodontitis and several non-communicable diseases including cardiovascular disease, diabetes, chronic kidney disease, respiratory diseases, and cognitive disorders [3,4]. Being the sixth most common human disease, severe periodontitis affects more than one tenth of the global adult population, hence representing a substantial health and socio-economic burden[5]. Notwithstanding the fact that inflammation is initiated by bacteria residing in the biofilm at gingival or oral mucosal surface, intrinsic host factors and environmental stressors determine the periodontal inflammatory response[6].

As the major constituents of periodontal soft tissue, gingiva are the first layer of defense against oral microorganisms [7]. Several types of cells are arrayed in an orderly manner in gingiva including epithelial cells (gingival, sulcular, junctional, and enamel), blast cells (fibroblast, cementoblasts, ameloblasts and osteoblasts), immune cells (neutrophils, T cells, B cells, monocytes, macrophage and dendritic cells), endothelial cells and neural cells, and undifferentiated mesenchymal cells [8-10]. Emerging evidence has indicated that structural cells, such as fibroblast, as well as T cells and macrophages in gingiva, are highly heterogeneous[11-13], leading, to various degrees, to periodontal inflammatory response. Nowadays, the diversity of cell subsets has been recognized as a substrate for new therapeutic strategies. For instance, the Treg-recruiting formulation system has been injected to treat severe experimental periodontitis[14]. Agents inducing the inflammatory-to-resolving conversion of macrophages were suggested to arrest periodontitis progression and stimulate bone regeneration[15]. However, the heterogeneity of in situ cells under periodontal health and disease conditions remain largely unknown.

On the other hand, studies have been performed to offer a global overview of molecular events of the inflammatory response in both *in vivo* and *in vitro* models. Multiple pathways were reported to be activated in in vitro infectious models, including inflammatory cytokines, reactive oxygen species generation, free radical scavenging, cell cycle, host defense mechanisms and Protein Kinase B/mitogen-activated protein kinase signaling pathway [16-21]. Studies of in vivo models, which focused on biopsy samples from periodontitis patients, have also demonstrated that genes in apoptosis, antimicrobial humoral response, antigen presentation, and regulation of metabolic process pathways were highly involved [22].Transcriptome analysis based on next-generation sequencing technology, especially RNA-seq, has revealed novel gene expression and splicing patterns in gingival biopsies from periodontitis patients compared with those from the healthy controls, such as upregulation of genes in the pathways of defense/immunity protein, receptor, protease, and signaling molecules, and downregulation in cytoskeletal and structural proteins [23]. Horie et al. performed transcriptome analysis of gingival fibroblast from periodontitis patients and identified osteogenic markers DLX5 and RUNX2 long variant as novel regulators [24]. Lundmark et al. provided the gene expression profiles of the periodontitis affected gingival biopsies, and found the inflamed area had upregulated expression of IGLL5, SSR4, MZB1 and XBP1, compared to the non-inflamed area [25].

Although transcriptome analysis based on bulk tissue RNA-seq has provided comprehensive gene expression profiles on paired gingival biopsies from both periodontitis-affected and healthy tissues, most studies only focused on the averaged transcriptional signatures on a preselected cell type or crossed all cell types in the whole tissue without the information on the cellular heterogeneity in the periodontal tissue. Herein, we reported the transcriptomic profiling of a total of 29,967 single cells of human gingival tissues from 2 patients with periodontitis and 2 healthy donors. We identified specific epithelial cell types associated with compromised epithelial defense in periodontitis. We detected endothelial cells highly associated with immune regulation and fibroblasts potentially involved in bone remodeling. We revealed the myeloid cells with diverse clusters dominated by pro-inflammatory phenotype. We further identified cell subtypes specifically expressing periodontitis related genes and investigated the cell-cell communication alteration between conditions. Overall, our study describes a transcriptome profile of human healthy and diseased gingival tissue at the single-cell resolution. It would offer a novel perspective on the periodontal inflammatory response and serve as a useful resource for the scientific community.

Results

scRNA-seq and cell typing of gingival tissues under healthy and disease conditions

To generate transcriptome profiles of human gingival tissues, samples from four donors were obtained. Two of the donors were diagnosed with periodontitis (P1 and P2). The other two samples were obtained from healthy volunteers who underwent crown lengthening procedures. Tissues were collected fresh, dissected, and digested into single cells (Figure 1A, Supplementary Table 1). For each sample, single cells were captured using the droplet-based microfluidic system Chromium (10X Genomics). The number of genes expressed differed in the various cell types, particularly between immune and non-immune cells (Supplementary Figure 1A and 1B).

Considering potential batch effects and background noises among samples, we applied SoupX[26] to correct ambient RNA in the background and merged these data using the CCA method in Seurat[27]. The single-cell data are presented in two-dimensional space using a uniform manifold approximation and projection (UMAP) method (Figure 1B). Clustering analysis cataloged these cells into ten distinct cell lineages annotated with canonical marker gene expression, thereby corresponding to epithelial cells, stromal cells (endothelial, vascular mural

and fibroblast), and immune cells (T, NK, B, plasma, myeloid and mast cells) (Figure 1C). Differential expressed genes (DEGs) of each cell type were computed and the top 5 DEGs were visualized (Supplementary Figure 1C).

Then we analyzed the distribution of the different cell lineages that accumulated among various conditions (Figure 1D). In healthy tissues, the epithelial cell has the largest proportion which comprises nearly 92% of all cells. However, in disease tissues, it only comprises 4.6%. In contrast, all the other cell types have their proportion increased in disease tissues, especially immune cells (T, B and plasma cells), which is consistent with the inflammation phenotype. Thus, although it was expected that immune activation, including the infiltration of the gingival tissue by T and B cells, would be involved in the pathogenesis, we provided a quantified description of the proportion change in the cell types. Our result also indicated that due to the enormous change in cell proportion, bulk RNA sequencing would be a less ideal method to measure gene expression changes when comparing healthy and periodontitis tissues, since the results would be largely confounded by cell proportion.

Next, we assessed the cell type-specific expression patterns of genes related to Mendelian disorders (based on OMIM database), which provided insights into the contribution of specific cell types to gingival abnormality (Figure 1E and Supplementary Table 2). Cell type-specific expression patterns confirmed fibroblast as particularly high expressors of the *C1S* and *C1R* gene, mutated in Ehlers-Danlos syndrome, periodontal type. Two reported gene mutants, *SOS1*[28] and *REST*[29], associated with gingival fibromatosis (GF), were highly expressed in endothelial cells. Additionally, the *CTSC* gene, mutations of which are responsible for aggressive periodontitis in juveniles, was found to be widely expressed in the non-immune cell population. While in immune cells, the expression of *CTSC* was highly enriched in myeloid cells.

Compromised epithelial barrier defense in periodontitis

The gingival epithelium, which consists of oral gingival epithelium, sulcular epithelium, and junctional epithelium, plays a vital role against bacterial invasion. We partitioned the epithelial cells into four diverse clusters based on previous reports[30,31]: junctional epithelium (JE1-JE2), basal (BAS1-BAS2), spinous (SPN1-SPN3), and differentiated granular keratinocytes (GRN1-GRN3) (Figure 2A, B). JE2 and JE1 dominated the epithelial cells in periodontitis, with significantly reduced GRN1/GRN2 and completely disappeared SPN1/SPN2 (Supplementary

Figure 2A). Immunohistochemical showed the outermost layer of the epithelium, referring to GRN and SPN, experienced severe ruptures in patients (Figure 2C). These data showed that the integrity of epithelial tissue is seriously damaged in periodontitis compared with the healthy controls.

We further analyzed the expression of characteristic genes and top 5 DEGs in each cluster (Figure 2D and Supplementary Figure 2B). For example, JE1 express amelogenesis associated proteins ODAM and ODAPH (C4orf26), which may participate in enamel mineralization. Meanwhile, we found both JE1 and JE2 expressed serum amyloid A family (SAA1 and SAA2), which can induce pathogenic Th17 and promote inflammation in experimental allergic encephalomyelitis and inflammatory bowel disease [32]. Besides, HLA-DRs (HLA-DRB1 and HLA-DRA) presenting extracellular pathogens were found highly expressed in JE2 and BAS1. BAS2 exhibited high levels of mitosis-related genes, denoting its possible stage at G2/m stages (Supplementary Figure 2C). The three spinous clusters were found to express *GSTA4* involved in anti-oxidative stress[33] and *CYB5A* detoxifying carcinogens from cigarette[34], thus playing essential roles in the detoxification of harmful foreign chemicals. GRN2 and GRN3 expressed a group of epidermal differentiation complex genes, which are essential for terminal differentiation of the human epidermis. Meanwhile, the two clusters expressed genes associated with the inflammatory response. We also noticed that GRN1-3 and JE1 expressed β-defensins (*DEFB1*) and *S100A7/8/9* against bacterial challenge, underscoring their role as the outermost layer in direct contact with the environment. Together, these data suggested that the damaged integrity of the epithelium barrier, leading to the loss of epithelial homeostasis-associated genes expressed by particular cell clusters in periodontitis, may constitute an important mechanism for periodontal pathogenicity. Analysis of the pathway in epithelial subsets was performed by gene set variation analysis (GSVA)[35] (Supplementary Figure 2D). The enriched terms highlight the difference in clusters and support the above inference about the function of each cell type, such as amelogenesis for JE1, DNA replication for BAS2, and keratinocyte differentiation for GRN3.

SCENIC[36] was utilized to identify different transcription factors (TFs) underlying the regulation of each epithelial phenotype (Figure 2E and Supplementary Table 3). For instance, Runx2 might be a potential regulator of the ODAM expression in junctional epithelium, thus playing a critical role in maintaining the integrity of the dentogingival junction[37,38]. Next, we used Monocle2[40,41] to further understand the relationship between epithelial cell states.

Combining trajectory plot and pseudotime analysis, we can form conclusions about the conditions and relationships among these clusters (Figure 2F). We found junctional epithelial and basal are both the beginning of the trajectory. This finding could be explained by the fact that junctional epithelial has the basement membrane-like structures with low differentiation and high regeneration ability[42,43]. Then, the trajectory experiences the spinous and end at the GRN3. We found 487 genes that vary significantly along pseudotime (Supplementary Table 4), and the heatmap showed the changes of expression pattern of 30 genes during the different stages (Figure 2G). The markers used for the subgroup definition, such as *FDCSP*, *KRT5,* and *SLURP1*, are indeed distributed along pseudotime trajectory. Some novel genes are specifically expressed in different stages, such as reversible epithelial-mesenchymal transition associated *CYR61* in JE[44] and epidermal differentiation associated *CALML5* in GRN[45]. This analysis may provide us deeper insights into the development of epithelial clusters that will be helpful in advancing periodontitis treatment and gingival tissue regeneration.

Endothelial cells highly associated with immune regulation and fibroblasts potentially involved in bone remodeling

To gain more insight into the heterogeneity of the gingival tissue, next, we focused on other non-immune clusters. The 1959 endothelial cells can be classified into three distinct subpopulations (Figure 3A). Endo_1 expressed *ACKR1*, *SELE*, and *SELP*. Endo_2 was characterized by *GJA4*, *HEY1*, and *NOTCH4*. We discovered a small group of lymphatic endothelial cells (LECs), named Endo_3, which expressed *PROX1*, *MMRN1*, and *CCL21*[46] (Figure 3B). The pathway analysis illustrates that Endo_1 is highly correlated with the immune response, response to interferon gamma, and upregulating adhesion molecules to tether or roll leukocyte. Meanwhile, Endo_1 is also involved in the regulation of blood pressure. In Endo_2, Notch signaling and other molecules contributing to endothelium development and migration were enriched. The enrichment pathways of Endo_3 meet the definition of LECs and are also involved in cell substrate adhesion (Figure 3C). Consistent with previous studies, we identified an epithelial cell state with high expression level of MHC class II genes such as *HLA-DRA*, *HLA-DRB1*, and *HLA-DPB1*, which was normally found in professional antigen-presenting cells[47]. The endothelial cell states in healthy tissue, particularly Endo_1, exhibited lower feature score of MHC class II compared with those in diseased tissues, suggesting the importance of

endothelial cells in gingival tissue-specific immunity (Figure 3D). Immunofluorescence assays for the MHC class II marker HLA-DR and the endothelial cell marker PECAM1 (CD31) further confirmed the existence of antigen-presenting endothelial cells in gingival tissues of periodontitis (Figure 3E).

Gingival fibroblasts have long been recognized as a heterogeneous population, but the extent of heterogeneity has hitherto remained poorly explored[48]. We identified two subclusters with multiple differentially expressed genes against each other (Figure 4A). Fibro_1 were characterized by high expression of *CXCL13*, *IL32*, and *SFRP2*, which are associated with B cell recruitment or pro-inflammatory[49,50]. Fibro_2 expressed higher levels of *OGN*, *PRELP*, and *RUNX2* compared with Fibro_1 (Figure 4B). Notably, forced OGN expression can increase the expression of RUNX2 and OCN, resulting in increased bone mass[51], while PRELP expressed in cartilage and basement membranes can impair osteoclastogenesis by inhibiting NF-kB[52]. Then we compared pathway enrichment between two fibroblast clusters (Figure 4C). Fibro_1 showed increased immune response pathways. Pathways associated with osteoblast development and bone remodeling are increased in Fibro_2. With immunofluorescent staining, we confirmed that OGN was detected in only a subset of fibroblast cells, which were labeled by anti-decorin staining. However, the role of Fibro_2 in bone remodeling needs to be further verified. Finally, we utilized scHCL[47] to verify the cluster identification and further explore the similarities between clusters (Figure 4E). By calculating Pearson correlation coefficients, the reliability of the cluster identification was confirmed. Interestingly, both epithelial and fibroblast cells were highly similar to those of esophageal origin, suggesting the effect of food intake on the cellular function of gingival fibroblast and epithelial cells.

Diverse immune cell subtypes with hyper-inflammatory response in periodontitis

Most of the tissue destruction in periodontitis is determined by the host immune response. To gain more insight into the host inflammatory and immune mechanisms in the gingival tissue, we performed sub-clustering on myeloid cell types containing 148 cells from healthy donors and 539 cells from patients. We revealed seven subtypes of myeloid cells[53] (Figure 5A). Three DC subsets were characterized by low expression of CD14, and three CD14-high expressing clusters were identified as macrophages based on their high expression of CD68, CD163, and MRC1[53,54] (Supplementary Figure 3A). Plasmacytoid DC (pDC), cDC1, and cDC2 were further

distinguished by specific expression of *LILRA4/GZMB/JCHAIN, BATF3/CLEC9A/CADM1*, *CD1C/CLEC10A/FCER1A*. Three clusters of CD14-high macrophages, Macro_PRDM1, Macro_NLRP3, and Macro_C1QA, were distinguished based on the expression of *FCGR2B/PRDM1/HES1*, *NLRP3/IL1B/EREG*, and *C1QA/SEPP1/SPOE*. There is a CD14+ monocyte cluster showing different features with DC and macrophages. All myeloid subtypes could be found in both normal and patient samples (Figure 5B). Differentially expressed marker genes between health and periodontitis were identified within each cell cluster (Figure 5C). The pDC group highly expressed Granzyme B, which has been reported to suppress T-cell expansion[55]. Expression of the pro-inflammatory cytokine gene *IL1B* was relatively high in Macro_PRDM1 and Macro_NLRP3. Further, Macro_NLRP3 highly expressing NLRP3 and S100A8 showed a strong pro-inflammatory phenotype. Macro_PRDM1 in periodontitis expressed a higher level of HES1, a transcription repressor controlling production of macrophage-derived chemokines in inflammatory arthritis [56] (Figure 5C).

The phenotypes of macrophages were analyzed in depth from angiogenesis and phagocytosis (Figure 5D). The gene associated with phagocytosis was highly expressed in Macro_C1QA, while Macro_NLRP3 dominated angiogenesis. Then we used RNA velocity[57] to study the developmental lineages of macrophages and projected the result onto the UMAP plot. With Macro_PRDM1 at the intermediate stage, macrophages in gingival tissue intend to differentiate from Macro_C1QA to pro-inflammatory Macro_NLRP3 phenotype. Besides, Macro_PRDM1 simultaneously resembled the signatures of M1 and M2 cells (Supplementary Figure 3B), indicating that the dichotomous concept of macrophage function was inappropriate[58]. Meanwhile, monocle2 was used to reveal the similar developmental lineages of macrophages, and the heatmap showed the top genes associated with latent time (Supplementary Figure 3C and 3D). Osteoclast differentiation-related protein CCL3 and CXCL2[59], but not other chemokine family genes, were found to be expressed in all macrophage subtypes in our data (Supplementary Figure 3E), suggesting their role in the regulation of macrophage-osteoclast differentiation.

Another diverse immune cell cluster is T and NK cells, which were divided into five subtypes (Figure 5E and 5G). CD4_CTLA4 highly express Treg-associated molecules, including *TNFRSF18*, *TNFRSF4*, *CTLA4*, and *FOXP3* (Supplementary Figure 3F). CD4_FOS from patients highly expressed immediate-early genes (*FOS*, *JUN*), which may be associated with T cell activation or the effect of enzymatic digestion[60]. Besides, CD8 and NK express cytotoxic

genes, including *GZMK*, *GZMA*, *GNLY*, and so on. Notably, a higher expression level of *CCR5* or CCR1 ligand *CCL4/CCL4L2/CCL3L3* by CD8 T cells was observed in patients, underscoring their role in inflammatory cell recruitment in periodontitis.

Differentially expressed genes and cell-cell interactions between periodontal health and disease conditions

Based on the previous bulk RNA sequencing[23,25,61-65], we noticed that most of the genes upregulated in periodontitis were found to be expressed in plasma and myeloid cells. By contrast, genes downregulated in periodontitis were expressed in epithelial cells (Figure 6A). Using microarrays or bulk RNA sequencing reflected an average of expression profiles of all cells on whole biopsies, largely confounded by the enormous change in cell proportion; thus, differential expression analysis within individual cell types was performed (Figure 6B). In endothelial cells, HLA class II molecules *HLA-DRB5* and *CLEC3B* associated with extracellular proteolysis were upregulated in periodontitis. In fibroblast, a metalloprotease *ADAM12* and *CFB* involved in activating B cells was increased in disease. The expression level of two C-C motif chemokine ligands, *CCL4L2* and *CCL3L3*, was elevated in myeloid from patients. We also found *PDCD1* was highly upregulated in the CD4_CTLA4 cluster from periodontitis, which indicates that PD-1 pathway may contribute to the protective effect of Treg in disease stage[66]. *CTSW* related to cytotoxic capacity was found to be upregulated in CD8_GZMK in periodontitis. These genes, previously masked in the mean expression data, provided novel insights for the characterization of periodontitis and will be helpful in advancing its therapy.

Periodontitis is a process of inflammation in the gingival tissue, which involves the cross talk of multiple cell types. Here, we used CellPhoneDB[67] to profile the communication among cell types in healthy and periodontitis tissues (Figure 6C). In both healthy and periodontitis tissues, the interactions between epithelial, endothelial, and fibroblast cells are the strongest, while others are relatively weak. Generally, the overall cell-cell interaction increased in the periodontitis tissue. We showed the top 10 cell-cell interactions increased in patients (Figure 6D). Myeloid and B cells interact more frequently with other cells. Further, significant differences in the relative contribution of chemokine receptor-ligand pairs could be observed between healthy and diseased gingiva (Figure 6E). Notably, the expression of CCR5 ligand

increased in periodontitis, which was consistent with our above findings. These pairs could play important roles in inflammation by recruiting inflammatory associated cells.

Discussion

Here, we present the first transcriptomic profiling of a total of 29,967 single cells of human gingival tissue using the scRNA-seq method. By identifying a comprehensive catalog of cell types and their phenotypes, revealing the altered gene expression profiling and cell-cell communication under diseased condition, our data highlight key areas for advances in the biology of periodontitis that will be helpful in the diagnosis and treatment of periodontitis.

Firstly, the epithelial is more heterogeneous than hitherto appreciated. While previous studies had indicated that a compromised epithelial barrier defense, enabling easier penetration of toxins or bacteria into the connective tissue, constituted an important mechanism for periodontitis, our data further identified specific cell clusters with decreased expression of epithelial homeostasis-associated genes, that contributed to pathogenicity of periodontal disease. Despite of the considerable efforts of previous researchers, molecular regulations of gingival development remained elusive. Using single-cell sequencing, different TFs underlying the regulation of individual epithelial phenotype were identified, including some novel molecules which had never been reported in this field before. For example, *POU2F2* was highly expressed in SPN2. The role of *POU2F2* in B cell and epithelial-derived cancer cells has been elucidated, but studies related to gingival development are still lacking.

Moreover, we detected 3 types of endothelial cells, 2 types of fibroblasts and different types of immune cells. We show the presence of the recently identified antigen-presenting endothelial cells in periodontitis[47], as revealed by the combined expression of CD31 and MHC class II genes. We identified a fibroblast subset potentially involved in bone remodeling. Heterogeneity of myeloid cells was depicted, and novel phenotypes of macrophage, hitherto considered dichotomous, were revealed. Our analysis confirmed that the in vitro characterized M1 and M2 cells do not reproduce the given tissue featured with a distinct local environment[68]. More samples are needed to study the function of different DC subgroups in periodontal and the differentiation of macrophages under different disease states. CCR5-ligand upregulated in cytotoxic CD8 T cells, might imply a potential treatment option by blocking cell recruitment. Besides, some genes associated with pro-inflammatory reactions were found to be expressed in

the clinically healthy individuals, suggesting the complicated mechanisms involved in the oral immune system to maintain immunological homeostasis.

Finally, we further investigated the biology of periodontitis in three directions: comparison between healthy and periodontitis group; profiles of known periodontitis related genes; and the cell-cell communication alteration between conditions. Our single-cell profiles not only provide an abundance of resources on cell types and interaction networks in the human gingiva, but also offer insights into the biological foundation of periodontal homeostasis and diseases, which potentially serve as the basis of therapeutic options.

In gingiva, an important soft tissue within the periodontium, the number of epithelial cells is far more than the number of immune cells. Under pathological conditions, cell types are not affected equally in the development of periodontitis. Thus, scRNA-seq becomes the most impartial and effective approach to obtain the transcriptome of each cell type in the gingiva. Admittedly, the complexity of the periodontitis could not be fully grasped, as the gingiva was located in restricted areas and the number of samples that were sequenced was also limited. Nevertheless, given the robustness of scRNA-seq, it is possible to scale up the current study to provide much improved resolution in the future.

Material and Methods

Sample collection and ethics approval

Collection of samples was approved by the Ethics Committee of Shanghai 9th People's Hospital in China (SH9H-2019-T158-2). The experiments conformed to the principles of the Helsinki Declaration revised in 2008. All donors gave informed consent prior to participation into the study. Clinical assessment and biopsy sampling were conducted at the Department of Oral Implantology, Shanghai 9th People's Hospital, Shanghai Jiaotong University, China. For scRNA-seq, periodontal tissues of two healthy individuals and two patients with periodontitis were collected. For inclusion in this study, patients with Stage III or IV periodontitis were diagnosed according to the new classification of periodontitis[69]. Samples were collected from healthy volunteers during crown lengthening procedure and from patients with periodontitis during open flap debridement procedure. The demographic information and clinical parameters of the two groups are shown in Supplementary Table 1.

Immunofluorescence

Human periodontal soft tissue biopsies were incubated with primary antibodies diluted in 3% BSA/PBS overnight at 4℃. Primary antibodies used include Decorin (1:50, ab175404, Abcam), Osteoglycin (1:50, sc-374463, Santa Cruz), HLA-DR (1:50, ab92511, Abcam), and CD31 (1:50, ab9498, Abcam). Next day samples were incubated in Alexa-fluor 488, 594 Goat anti-Mouse or Goat anti-Rabbit secondary antibodies (Jackson Immunoresearch). Nuclei were counterstained with DAPI. Pictures were acquired using Zeiss LSM 880.

Preparation of single-cell suspensions

Once the sample was retrieved, it was dissociated and processed for scRNA-seq immediately. Periodontal soft tissue samples were minced into small fragments of less than 1 mm3 by surgical scissors and dissociated into single cells in dissociation solution (2 mg/mL IV collagenase, 2 Unit/mL Dispase II in $Ca2^{+}$- and $Mg2^{+}$-free HBSS) covered with tinfoil on a shaker (shaking speed of 200 rpm) at 37℃ for 60 min. 0.1 μg/mL DNase I was added in the last 10 min. The dissociated tissue was filtered to ensure single-cell suspension using 100μm, 40μm cell strainers (Falcon) successively. Cells were subjected to red blood cell lysis for 10 min and centrifuged and resuspended (500g, 10min) twice. After resuspension in defined volumes of PBS + 0.4% BSA, 10μL of the cell suspension was used for cell counting by an automated cell counter (Thermofisher) to determine the concentration of live cells. Single-cell samples with final cell viability above 90% and final concentration of 600-1200 cells μL^{-1} were stored on ice until further processing. The whole procedure was performed on ice whenever possible.

Single-cell RNA sequencing and read processing

Single cells from independent periodontal samples were captured in four batches using the 10X Chromium system (10X Genomics). The cells were partitioned into Gel Bead-In-Emulsions and barcoded cDNA libraries, then prepared using the Chromium Single Cell 3' library & Gel Bead Kit v3 (10X Genomics). Single-cell libraries were sequenced in 100 bp paired-end configuration using an Illumina NovaSeq and mapped to the GRCh38 human reference genome using the Cell Ranger toolkit (version 3.0.0). The preliminary data analysis generated a file containing a barcodes table, a genes table, and a gene expression matrix. Next, we obtained an overview

website containing a considerable amount of information, such as number of cells, median number of detected genes, sequencing saturation, and sequencing depth.

Filtering and normalization of scRNA-seq data

We installed R (version 3.5.1) and Seurat R package (version 3.1.5) for downstream analysis. First, the substantial background levels of ambient RNA in the single-cell suspension caused problems for subsequent analysis. Thus, we applied SoupX (version 1.4.5) for background correction. Next, for quality control of each matrix, lowly detected genes (<0.1% cells) and cells with a small number of genes (<350 genes) were discarded from the downstream analysis. Then, we filtered out unhealthy cells that generally have high mitochondrial mRNA loads (>20%) and high ribosome RNA loads (>40%). We found that different cell types expressed different numbers of genes, particularly between immune and non-immune cells. Thus, we applied a slightly different criteria to remove supposed *ambient RNA contamination* (<1700 UMI for healthy detected per cell) and potential double droplets (>6500 genes for healthy, and >4000 genes for periodontitis tissues detected per cell). After the step above, we obtained 10,501 high quality periodontitis cells (median UMI: 3544; median 1203 genes/cell) and 19,977 healthy periodontal cells (median UMI: 16,455; median 3169 genes/cell).

Dimension reduction, unsupervised clustering, annotation, and visualization

The expression value of each gene was first normalized by TPM/10 and then log-transformed (NormalizeData function in Seurat with default parameters). Using the variation stabilizing transformation (vst) method, the top 2000 variable genes were selected in each matrix and were used as input for the 'FindIntegrationAnchors' function. The four expression matrices were then integrated with the 'IntegrateData' function. The integrated data were dimension reduced with principal component analysis (PCA; top 30 dimensions) first and then further reduced to two dimensions with UMAP which was also used to visualize the clusters. The nearest neighbors were defined among cells with KNN method (FindNeighbors), and cells were then grouped with Louvain algorithm (FindClusters in Seurat, resolution equal to 1.5; PCA:top 25 dimensions). Specifically, we removed one cluster, considered as contamination, due to the co-expressed markers of plasma and epithelial cells (data not shown). Finally, we retained a total of 29,967 cells after stringent quality controls for further analysis, with 19,806 (66.09%) from healthy

tissues and 10,161 from the periodontitis tissues. Annotation of the clusters was performed by checking known markers for cell types that potentially would exist in the sample, and some clusters were merged as they were annotated as a major cell type. Average expression levels of each subtype marker was calculated by AddModuleScore in Seurat with default parameters. For sub-clustering, cells from a major cell type were taken as the input. We performed dimension reduction, clustering, and annotation using the same method as described above.

Detection of differentially expressed genes (DEGs)

To obtain the DEG list of each cluster, only the genes expressed in more than 30% of that cluster were considered, and the expression in all other cells was used as background. For statistical test, we used the default Wilcoxon test implemented in Seurat. DEGs were defined as genes whose log fold-change was over 0.2 compared to the background, and with a q-value (FDR) smaller than 0.05.

Gene set enrichment analysis

We conducted the gene set enrichment analysis for DEGs of each cluster using clusterProfiler[70], GSVA, and GSEABase packages, with which the enriched GO biological process terms were calculated.

SCENIC analysis

SCENIC analysis was conducted as described previously.[36] We used the pySCENIC package (version 0.10.3), a lightning-fast python implementation of the SCENIC pipeline. The differentially activated TFs of each subcluster were identified by the Wilcoxon rank sum test against all the other cells of the same cell type.

Reconstructing gingival epithelial differentiation trajectories using Monocle2

Cell fate decisions and pseudotime trajectories were reconstructed using the Monocle2 R package (version 2.10.1). Firstly, these Seurat data which included eight types of epithelial were imported into Monocle2. Genes that were expressed in at least ten cells were used, and only genes expressed in .3% of cells were kept. We used thresholds on the cells local density (r) and nearest distance (d) to determine the number of clusters. We used ‘dpFeature’ to find the

differential genes (mean_expression >= 0.3 & dispersion_empirical >= 1 * dispersion_fit). These DEGs, which were a set of ordering genes, were then used to perform the dimension reduction and the trajectory analysis. Once we established a trajectory, we used the differential Gene Test function to find genes that had an expression pattern that varied according to pseudotime.

RNA velocity analysis

RNA velocity analysis was performed using the scVelo (version 0.2.1). Briefly, spliced/unspliced reads were annotated by velocyto.py with CellRanger (version 3.0.0), generating BAM files and then saved in .loom files. The .loom files were then loaded to python (version 3.8.3) using the anndata.read_loom function to generate count tables for splicing and unsplicing reads. Then we used dynamical model from scVelo to calculate. Lastly, the velocity vector arrows were projected onto the UMAP plot which was obtained in Seurat.

Profiling the cell-cell communication in healthy and periodontitis samples

The cell-cell communication was measured by quantification of ligand-receptor pairs among different cell types. Gene expression matrices and metadata with major cell annotations were used as input for the CellphoneDB or CellChat[71] software. The default CellPhoneDB database and parameters were used. Healthy and periodontitis data was computed separately. The cell-cell network was visualized with circlize (Version 0.4.10).

Author contributions

H.L. and B.L. conceived the project. S.Q., Q.H., Y.Z., and L.Z. participated in the data analysis. J.S. advised the data analysis. S.Q., L.Z, and J.S. collected the human donor gingiva, performed the phenotyping and dissection. Q.H. and L.Z. prepared the nuclei sample and performed the single-nuclei RNA-seq. S.Q., Q.H., H.L. and B.L. wrote the manuscript with input from all other authors. All authors proofread the manuscript.

Competing interests

The authors declare no competing interests.

References

1 Nunez, J., Vignoletti, F., Caffesse, R. G. & Sanz, M. Cellular therapy in periodontal regeneration. *Periodontol 2000* **79**, 107-116, doi:10.1111/prd.12250 (2019).

2 Ouchi, T. & Nakagawa, T. Mesenchymal stem cell-based tissue regeneration therapies for periodontitis. *Regen Ther* **14**, 72-78, doi:10.1016/j.reth.2019.12.011 (2020).

3 Tonetti, M. S., Van Dyke, T. E. & working group 1 of the joint, E. F. P. A. A. P. w. Periodontitis and atherosclerotic cardiovascular disease: consensus report of the Joint EFP/AAPWorkshop on Periodontitis and Systemic Diseases. *J Periodontol* **84 Suppl 4S**, S24-S29, doi:10.1902/jop.2013.1340019 (2013).

4 Genco, R. J. & Sanz, M. Clinical and public health implications of periodontal and systemic diseases: An overview. *Periodontology 2000* **83**, 7-13, doi:10.1111/prd.12344 (2020).

5 Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386**, 743-800, doi:10.1016/s0140-6736(15)60692-4 (2015).

6 Bartold, P. M. & Van Dyke, T. E. Periodontitis: a host-mediated disruption of microbial homeostasis. Unlearning learned concepts. *Periodontol 2000* **62**, 203-217, doi:10.1111/j.1600-0757.2012.00450.x (2013).

7 Schroeder, H. E. & Listgarten, M. A. The gingival tissues: the architecture of periodontal protection. *Periodontol 2000* **13**, 91-120, doi:10.1111/j.1600-0757.1997.tb00097.x (1997).

8 Donos, N. The periodontal pocket. *Periodontol 2000* **76**, 7-15, doi:10.1111/prd.12203 (2018).

9 Nunez, J., Vignoletti, F., Caffesse, R. G. & Sanz, M. Cellular therapy in periodontal regeneration. *Periodontology 2000* **79**, 107-116, doi:10.1111/prd.12250 (2019).

10 Moutsopoulos, N. M. & Konkel, J. E. Tissue-Specific Immunity at the Oral Mucosal Barrier. *Trends Immunol* **39**, 276-287, doi:10.1016/j.it.2017.08.005 (2018).

11 Parisi, L. *et al.* Macrophage Polarization in Chronic Inflammatory Diseases: Killers or Builders? *J Immunol Res* **2018**, 8917804, doi:10.1155/2018/8917804 (2018).

12 Smith, P. C., Martínez, C., Martínez, J. & Mcculloch, C. A. Role of Fibroblast Populations in Periodontal Wound Healing and Tissue Remodeling. *Front. Physiol.* **10**, 270- (2019).

13 Alvarez, C. *et al.* Regulatory T Lymphocytes in Periodontitis: A Translational View. *Mediators Inflamm.* **2018**, 7806912, doi:10.1155/2018/7806912 (2018).

14 Glowacki, A. J. *et al.* Prevention of inflammation-mediated bone loss in murine and canine periodontal disease via recruitment of regulatory lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18525-18530, doi:10.1073/pnas.1302829110 (2013).

15 Sima, C., Viniegra, A. & Glogauer, M. Macrophage immunomodulation in chronic osteolytic diseases—the case of periodontitis. *Journal of Leukocyte Biology* **105**, 473-487, doi:10.1002/jlb.1ru0818-310r (2019).

16 Xie, Y., Sun, M., Xia, Y. & Shu, R. An RNA-seq screen of P. gingivalis LPS treated human gingival fibroblasts. *Arch Oral Biol* **88**, 77-84, doi:10.1016/j.archoralbio.2018.01.002 (2018).

17 Wang, P. L., Azuma, Y., Shinohara, M. & Ohura, K. Toll-like receptor 4-mediated signal pathway induced by Porphyromonas gingivalis lipopolysaccharide in human gingival fibroblasts. *Biochem Bioph Res Co* **273**, 1161-1167, doi:10.1006/bbrc.2000.3060 (2000).

18 Wu, X., Zhang, G., Feng, X., Li, P. & Tan, Y. Transcriptome analysis of human periodontal ligament fibroblasts exposed to Porphyromonas gingivalis LPS. *Arch Oral Biol* **110**, 104632, doi:10.1016/j.archoralbio.2019.104632 (2020).

19 Ahn, S. H. *et al.* Transcriptome profiling analysis of senescent gingival fibroblasts in response to Fusobacterium nucleatum infection. *PLoS One* **12**, e0188755, doi:10.1371/journal.pone.0188755 (2017).

20 Kang, W. *et al.* Time-Course Transcriptome Analysis for Drug Repositioning in Fusobacterium nucleatum-Infected Human Gingival Fibroblasts. *Front Cell Dev Biol* **7**, 204, doi:10.3389/fcell.2019.00204 (2019).

21 Kang, W., Sun, T., Tang, D., Zhou, J. & Feng, Q. Time-Course Transcriptome Analysis of Gingiva-Derived Mesenchymal Stem Cells Reveals That Fusobacterium nucleatum Triggers Oncogene Expression in the Process of Cell Differentiation. *Front Cell Dev Biol* **7**, 359, doi:10.3389/fcell.2019.00359 (2019).

22 Demmer, R. T. *et al.* Transcriptomes in Healthy and Diseased Gingival Tissues. *Journal of Periodontology* **79**, 2112-2124, doi:10.1902/jop.2008.080139 (2008).
23 Kim, Y. G. *et al.* Transcriptome sequencing of gingival biopsies from chronic periodontitis patients reveals novel gene expression and splicing patterns. *Hum Genomics* **10**, 28, doi:10.1186/s40246-016-0084-0 (2016).
24 Horie, M. *et al.* Transcriptome analysis of periodontitis-associated fibroblasts by CAGE sequencing identified DLX5 and RUNX2 long variant as novel regulators involved in periodontitis. *Sci Rep* **6**, 33666, doi:10.1038/srep33666 (2016).
25 Lundmark, A. *et al.* Gene expression profiling of periodontitis-affected gingival tissue by spatial transcriptomics. *Sci Rep* **8**, 9370, doi:10.1038/s41598-018-27627-3 (2018).
26 Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet based single-cell RNA sequencing data. 303727, doi:10.1101/303727 %J bioRxiv (2020).
27 Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411-420, doi:10.1038/nbt.4096 (2018).
28 Hart, T. C. *et al.* A mutation in the SOS1 gene causes hereditary gingival fibromatosis type 1. *Am J Hum Genet* **70**, 943-954, doi:10.1086/339689 (2002).
29 Bayram, Y. *et al.* REST Final-Exon-Truncating Mutations Cause Hereditary Gingival Fibromatosis. *Am J Hum Genet* **101**, 149-156, doi:10.1016/j.ajhg.2017.06.006 (2017).
30 Dabija-Wolter, G., Bakken, V., Cimpan, M. R., Johannessen, A. C. & Costea, D. E. In vitro reconstruction of human junctional and sulcular epithelium. *J Oral Pathol Med* **42**, 396-404, doi:10.1111/jop.12005 (2013).
31 Cheng, J. B. *et al.* Transcriptional Programming of Normal and Inflamed Human Epidermis at Single-Cell Resolution. *Cell Rep* **25**, 871-883, doi:10.1016/j.celrep.2018.09.006 (2018).
32 Lee, J. Y. *et al.* Serum Amyloid A Proteins Induce Pathogenic Th17 Cells and Promote Inflammatory Disease. *Cell* **180**, 79-91 e16, doi:10.1016/j.cell.2019.11.026 (2020).
33 Singhal, S. S. *et al.* Antioxidant role of glutathione S-transferases: 4-Hydroxynonenal, a key molecule in stress-mediated signaling. *Toxicol Appl Pharmacol* **289**, 361-370, doi:10.1016/j.taap.2015.10.006 (2015).

34 Blanke, K. L. *et al.* Polymorphisms in the carcinogen detoxification genes CYB5A and CYB5R3 and breast cancer risk in African American women. *Cancer Causes Control* **25**, 1513-1521, doi:10.1007/s10552-014-0454-7 (2014).

35 Hanzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7, doi:10.1186/1471-2105-14-7 (2013).

36 Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* **14**, 1083-1086, doi:10.1038/nmeth.4463 (2017).

37 Xu, C. *et al.* Epithelium-specific Runx2 Knockout Mice Display Junctional Epithelium and Alveolar Bone Defects. *Oral Dis*, doi:10.1111/odi.13647 (2020).

38 Lee, H. K. *et al.* The odontogenic ameloblast-associated protein (ODAM) cooperates with RUNX2 and modulates enamel mineralization via regulation of MMP-20. *J Cell Biochem* **111**, 755-767, doi:10.1002/jcb.22766 (2010).

39 Kawasaki, K. *et al.* Expression of Sox genes in tooth development. *Int J Dev Biol* **59**, 471-478, doi:10.1387/ijdb.150192ao (2015).

40 Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381-386, doi:10.1038/nbt.2859 (2014).

41 Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* **14**, 979-982, doi:10.1038/nmeth.4402 (2017).

42 Jiang, Q., Yu, Y., Ruan, H., Luo, Y. & Guo, X. Morphological and functional characteristics of human gingival junctional epithelium. *BMC Oral Health* **14**, 30, doi:10.1186/1472-6831-14-30 (2014).

43 Nakamura, M. Histological and immunological characteristics of the junctional epithelium. *Jpn Dent Sci Rev* **54**, 59-65, doi:10.1016/j.jdsr.2017.11.004 (2018).

44 Chai, J. *et al.* CCN1 induces a reversible epithelial-mesenchymal transition in gastric epithelial cells. *Lab Invest* **90**, 1140-1151, doi:10.1038/labinvest.2010.101 (2010).

45 Sun, B. K. *et al.* CALML5 is a ZNF750- and TINCR-induced protein that binds stratifin to regulate epidermal differentiation. *Genes Dev* **29**, 2225-2230, doi:10.1101/gad.267708.115 (2015).

46 Fujimoto, N. *et al.* Single-cell mapping reveals new markers and functions of lymphatic endothelial cells in lymph nodes. *PLoS Biol* **18**, e3000704, doi:10.1371/journal.pbio.3000704 (2020).

47 Han, X. *et al.* Construction of a human cell landscape at single-cell level. *Nature* **581**, 303-309, doi:10.1038/s41586-020-2157-4 (2020).

48 Hinz, B. Matrix mechanics and regulation of the fibroblast phenotype. *Periodontology 2000* **63**, 14-28, doi:10.1111/prd.12030 (2013).

49 Armas-Gonzalez, E. *et al.* Role of CXCL13 and CCL20 in the recruitment of B cells to inflammatory foci in chronic arthritis. *Arthritis Res Ther* **20**, 114, doi:10.1186/s13075-018-1611-2 (2018).

50 Kim, S. H., Han, S. Y., Azam, T., Yoon, D. Y. & Dinarello, C. A. Interleukin-32: a cytokine and inducer of TNFalpha. *Immunity* **22**, 131-142, doi:10.1016/j.immuni.2004.12.003 (2005).

51 Chen, X. *et al.* Effects of Osteoglycin (OGN) on treating senile osteoporosis by regulating MSCs. *BMC Musculoskelet Disord* **18**, 423, doi:10.1186/s12891-017-1779-7 (2017).

52 Rucci, N. *et al.* The glycosaminoglycan-binding domain of PRELP acts as a cell type-specific NF-kappaB inhibitor that impairs osteoclastogenesis. *J Cell Biol* **187**, 669-683, doi:10.1083/jcb.200906014 (2009).

53 Zhang, L. *et al.* Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer. *Cell* **181**, 442-459 e429, doi:10.1016/j.cell.2020.03.048 (2020).

54 Tang-Huau, T. L. *et al.* Human in vivo-generated monocyte-derived dendritic cells and macrophages cross-present antigens through a vacuolar pathway. *Nat Commun* **9**, 2570, doi:10.1038/s41467-018-04985-0 (2018).

55 Jahrsdorfer, B. *et al.* Granzyme B produced by human plasmacytoid dendritic cells suppresses T-cell expansion. *Blood* **115**, 1156-1165, doi:10.1182/blood-2009-07-235382 (2010).

56 Shang, Y. *et al.* The transcriptional repressor Hes1 attenuates inflammation by regulating transcription elongation. *Nature Immunology* (2016).

57 Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol*, doi:10.1038/s41587-020-0591-3 (2020).

58 Murray, P. J. *et al.* Macrophage activation and polarization: nomenclature and experimental guidelines. *Immunity* **41**, 14-20, doi:10.1016/j.immuni.2014.06.008 (2014).

59 Dapunt, U., Maurer, S., Giese, T., Gaida, M. M. & Hansch, G. M. The macrophage inflammatory proteins MIP1alpha (CCL3) and MIP2alpha (CXCL2) in implant-associated osteomyelitis: linking inflammation to bone degradation. *Mediators Inflamm* **2014**, 728619, doi:10.1155/2014/728619 (2014).

60 van den Brink, S. C. *et al.* Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat Methods* **14**, 935-936, doi:10.1038/nmeth.4437 (2017).

61 Davanian, H. *et al.* Gene expression profiles in paired gingival biopsies from periodontitis-affected and healthy tissues revealed by massively parallel sequencing. *PLoS One* **7**, e46440, doi:10.1371/journal.pone.0046440 (2012).

62 Demmer, R. T. *et al.* Transcriptomes in healthy and diseased gingival tissues. *J Periodontol* **79**, 2112-2124, doi:10.1902/jop.2008.080139 (2008).

63 Beikler, T., Peters, U., Prior, K., Eisenacher, M. & Flemmig, T. F. Gene expression in periodontal tissues following treatment. *BMC Med Genomics* **1**, 30, doi:10.1186/1755-8794-1-30 (2008).

64 Becker, S. T. *et al.* Peri-implantitis versus periodontitis: functional differences indicated by transcriptome profiling. *Clin Implant Dent Relat Res* **16**, 401-411, doi:10.1111/cid.12001 (2014).

65 Lundmark, A. *et al.* Transcriptome analysis reveals mucin 4 to be highly associated with periodontitis and identifies pleckstrin as a link to systemic diseases. *Sci Rep* **5**, 18475, doi:10.1038/srep18475 (2015).

66 Wafula, P. O. *et al.* PD-1 but not CTLA-4 blockage abrogates the protective effect of regulatory T cells in a pregnancy murine model. *Am J Reprod Immunol* **62**, 283-292, doi:10.1111/j.1600-0897.2009.00737.x (2009).

67 Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* **15**, 1484-1506, doi:10.1038/s41596-020-0292-x (2020).

68 Amit, I., Winter, D. R. & Jung, S. The role of the local environment and epigenetics in shaping macrophage identity and their effect on tissue homeostasis. *Nature Immunology* **17**, 18-25, doi:10.1038/ni.3325 (2016).

69 Tonetti, M. S., Greenwell, H. & Kornman, K. S. Staging and grading of periodontitis: Framework and proposal of a new classification and case definition. *J Clin Periodontol* **45 Suppl 20**, S149-S161, doi:10.1111/jcpe.12945 (2018).

70 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287, doi:10.1089/omi.2011.0118 (2012).

71 Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. 2020.2007.2021.214387, doi:10.1101/2020.07.21.214387 %J bioRxiv (2020).

Figure 1 Overview of the clustering and annotation of the single-cell RNA sequencing data for gingival tissues.

A. Schematic depicting anatomical regions analyzed in this study (The graph was created with BioRender.com).
B. tSNE of the 29,967 cells profiled here, with each cell color-coded for (left to right): the corresponding patient and the associated cell type.
C. Well-known markers scores for annotation of gingival cell types visualized by violin plot.
D. Proportion of cell types from periodontitis and healthy conditions. (Nor1 and Nor2 are healthy donors; P1 and P2 are periodontitis patients)
E. Dot plot depicting gene expression levels and percentage of cells expressing genes associated with periodontal disease according to the OMIM database.

Figure 2 The heterogeneity and developmental trajectory of gingival epithelial.

A. Visualization of human gingival epithelial cells in disease and healthy conditions separately.
B. Markers for annotation of gingival epithelial cell types visualized by violin plot.
C. Immunohistochemical showed the outermost layer of the epithelium experienced severe ruptures in patients.
D. UMAP plots of epithelial cells, color-coded for functional molecules scores.
E. Heatmap showing the activity of TFs in each epithelial cell subtype. The TF activity is scored using scaled AUCell.
F. Monocle2-generated pseudotime trajectory of epithelial types imported from Seurat data, colored by cell-name designation (left) and colored in a gradient from dark to light yellow (right).
G. Heatmap for clustering the 30 genes that vary as a function of pseudotime. The 30 genes were divided into four clusters, representing the genes at the beginning stage, the transitory stage, and the end stage of developmental trajectory, respectively.

Figure 3 Three endothelial subclusters identified from the human gingival cells.

A. UMAP visualization of three endothelial subclusters from human gingival cells.
B. Violin plots showing the expression distribution of selected genes associated with functions in the endothelial clusters.
C. Top gene ontology term enrichment for endothelial clusters.
D. HLA-DRA showed expression specificity in the endothelial Endo_1. The gene expression levels are normalized and transformed as ln (CPM/10).
E. Immunofluorescent (IF) staining validation of endothelial subtypes. Red color showed the signal of HLA-DR staining; green color showed the signal of CD-31 staining (endothelial marker); and blue color showed DAPI staining.

Figure 4 Two fibroblast subclusters identified from the human gingival cells.

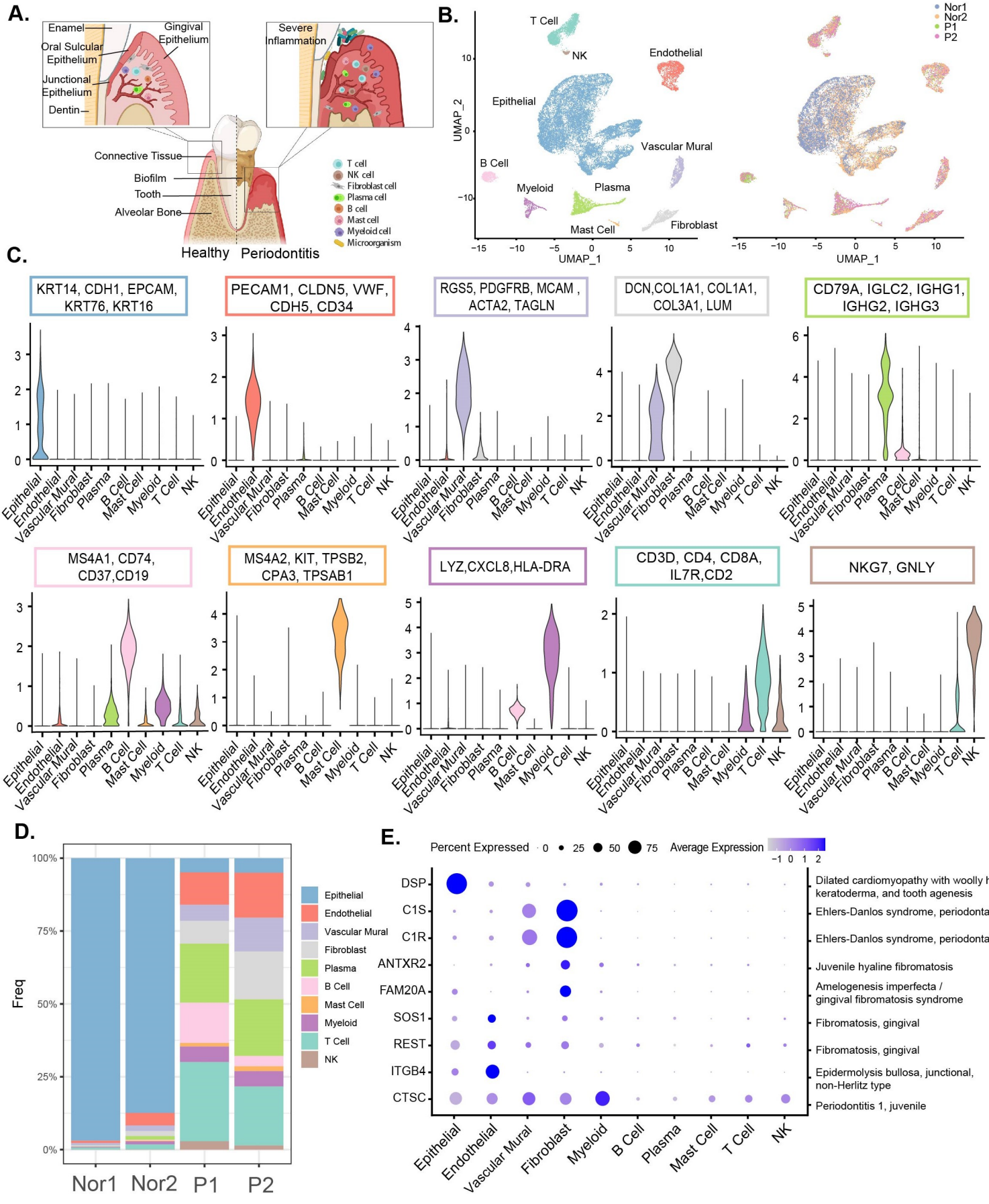
A. UMAP visualization of two fibroblast subclusters from human gingival cells.
B. Violin plots showing the expression distribution of selected genes associated with functions in the endothelial clusters. The gene expression levels are normalized and transformed as ln (CPM/10).
C. Differences in pathway activities scored per cell by GSVA between fibro_1 and fibro_2. Shown are *t* values from a linear model, corrected for fibro_1.
D. Immunofluorescent (IF) staining validation of fibroblast subtypes. The IF was performed on subepithelial region and connective tissue, respectively. Red color showed the signal of Decorin staining (fibroblast marker); green color showed the signal of Osteoglycin (OGN) staining; and blue color showed DAPI staining.
E. Application of scHCL analysis for non-immune cells. Each row represents one cell type in scHCL. Each column represents a cell cluster in our dataset. Pearson correlation coefficient was used to evaluate cell-type gene expression similarity. Red indicates a high correlation; blue indicates a low correlation.

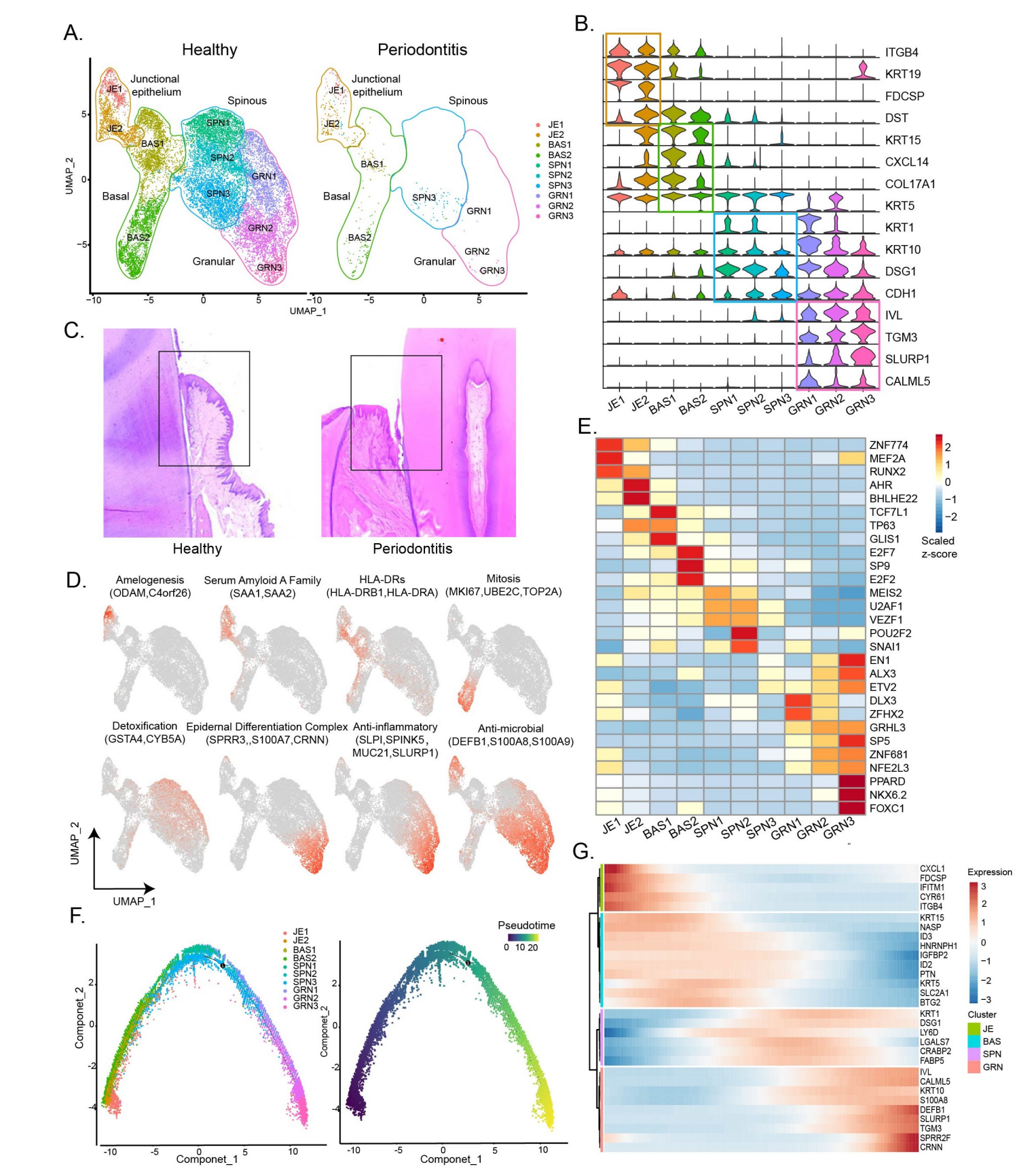
Figure 5 The heterogeneity of myeloid and T cells.

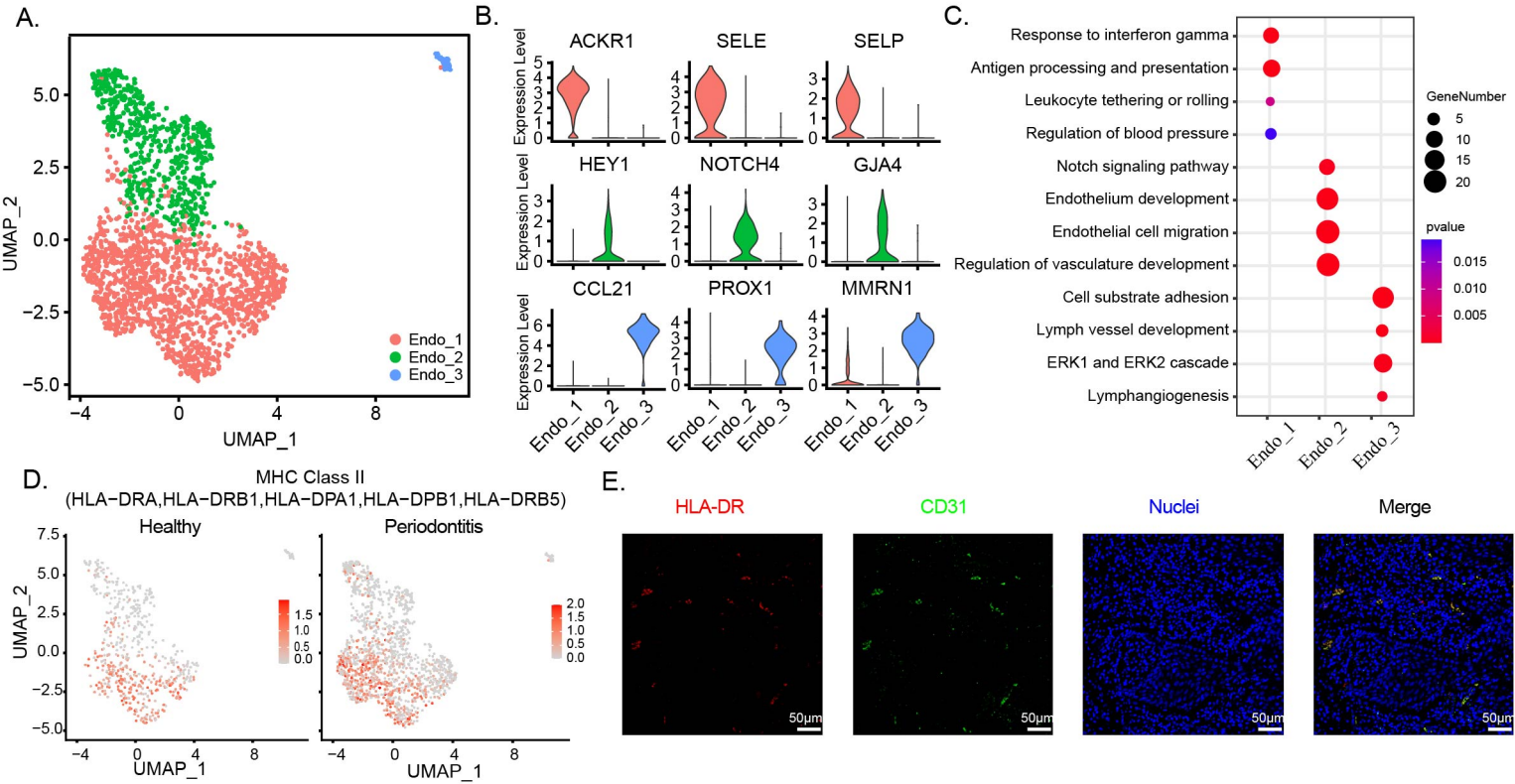
A. UMAP visualization of seven myeloid clusters from human gingival cells.
B. Relative proportion of myeloid cell subsets from different origin (healthy=2 and periodontitis=2).
C. Bubble heatmap showing marker genes across seven myeloid clusters from (A). Dot size indicates fraction of expressing cells, colored according to z-score normalized expression levels.
D. Violin plots showing the expression of angiogenesis- and phagocytosis-related genes in three macrophages clusters. The gene expression levels are normalized and transformed as ln (CPM/10).
E. Velocity field projected onto the UMAP plot of three macrophages (arrows represent average velocity).
F. UMAP visualization of seven myeloid clusters from human gingival cells.
G. Relative proportion of T cell subsets from different origin (healthy=2 and periodontitis=2).
H. Heatmap of T cell top 5 DEGs in each cell cluster in healthy and periodontitis.

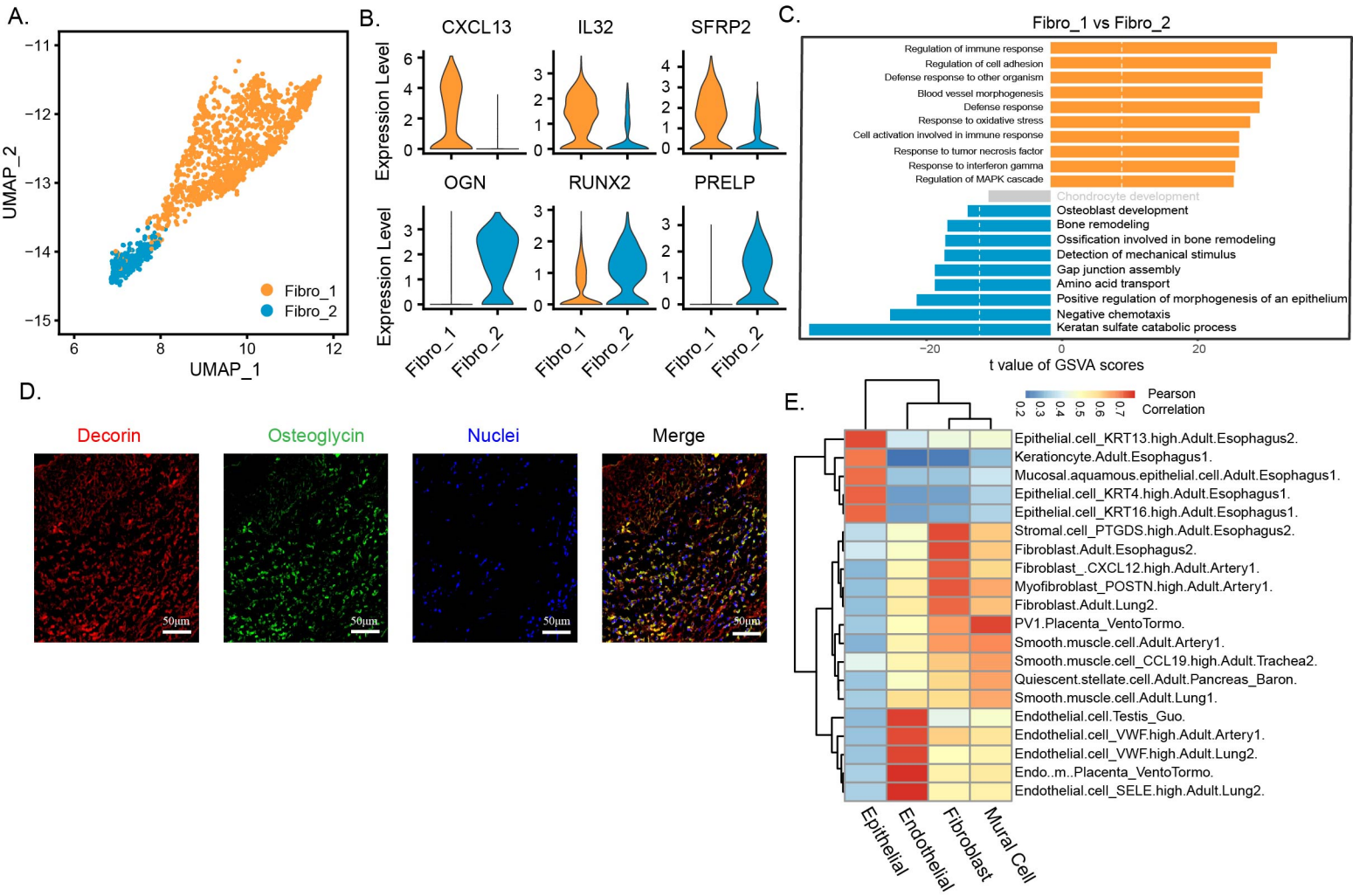
Figure 6 The difference among gingival cell types in healthy and periodontitis conditions.

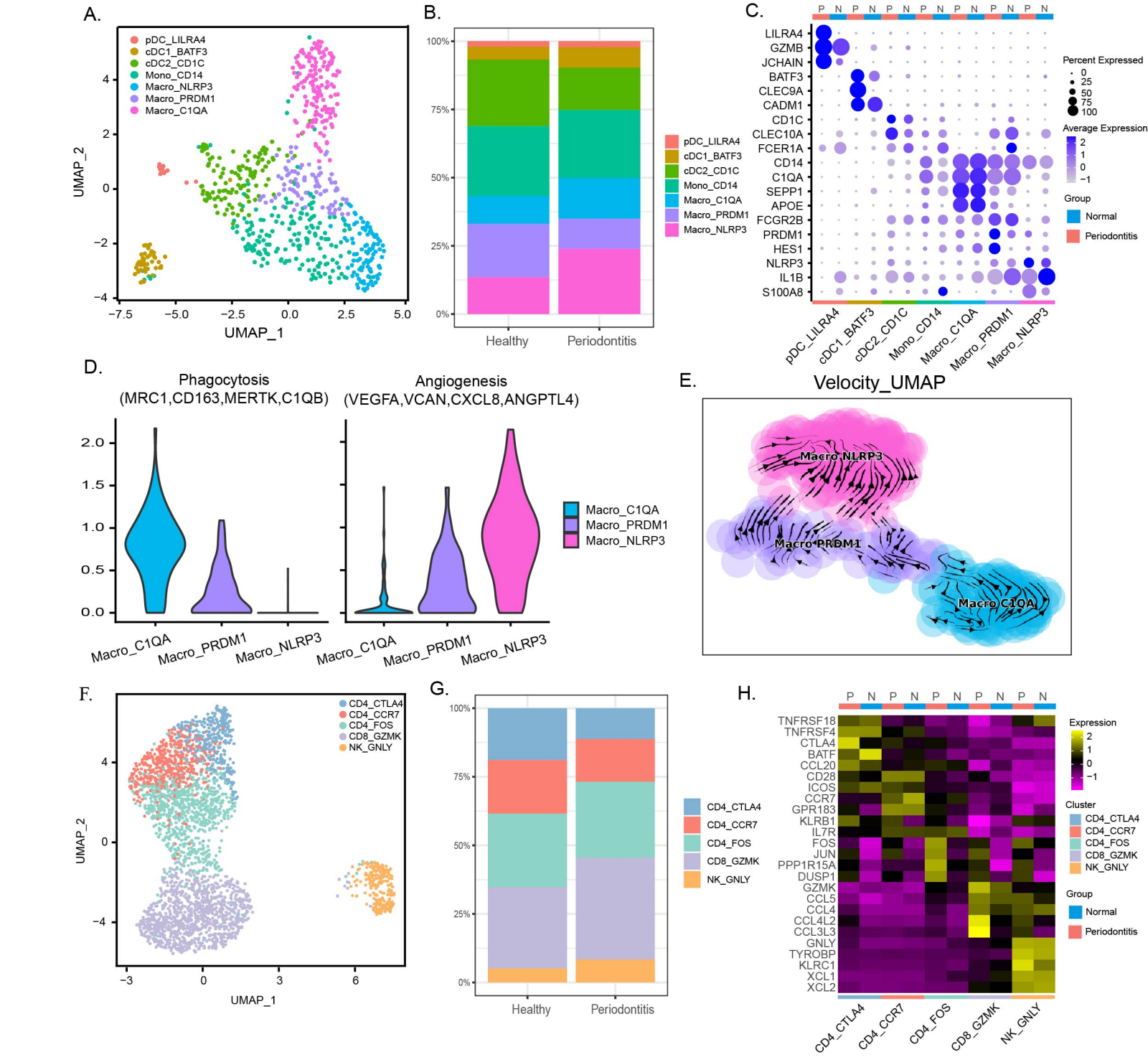
A. Bubble heatmap showing reported genes associated diseases clusters from (Figure 1B). Rose stands for reported increase in periodontitis, and turquoise stands for reported decrease in periodontitis. Dot size indicates fraction of expressing cells, colored according to z-score normalized expression levels.
B. Violin plots showing the expression of gene specific increasing in different clusters of periodontitis. The gene expression levels are normalized and transformed as ln (CPM/10).
C. Cell-cell interaction network in healthy and periodontitis condition. Colors and widths of edges represent number of interaction pairs between cell types.
D. The top 10 cell-cell interactions with the most significant increase in patients than healthy. The relative contribution of receptor-ligand in CCL pathway in normal and periodontitis. Yellow indicates that the common receptor-ligand is present in both the patient and the normal.

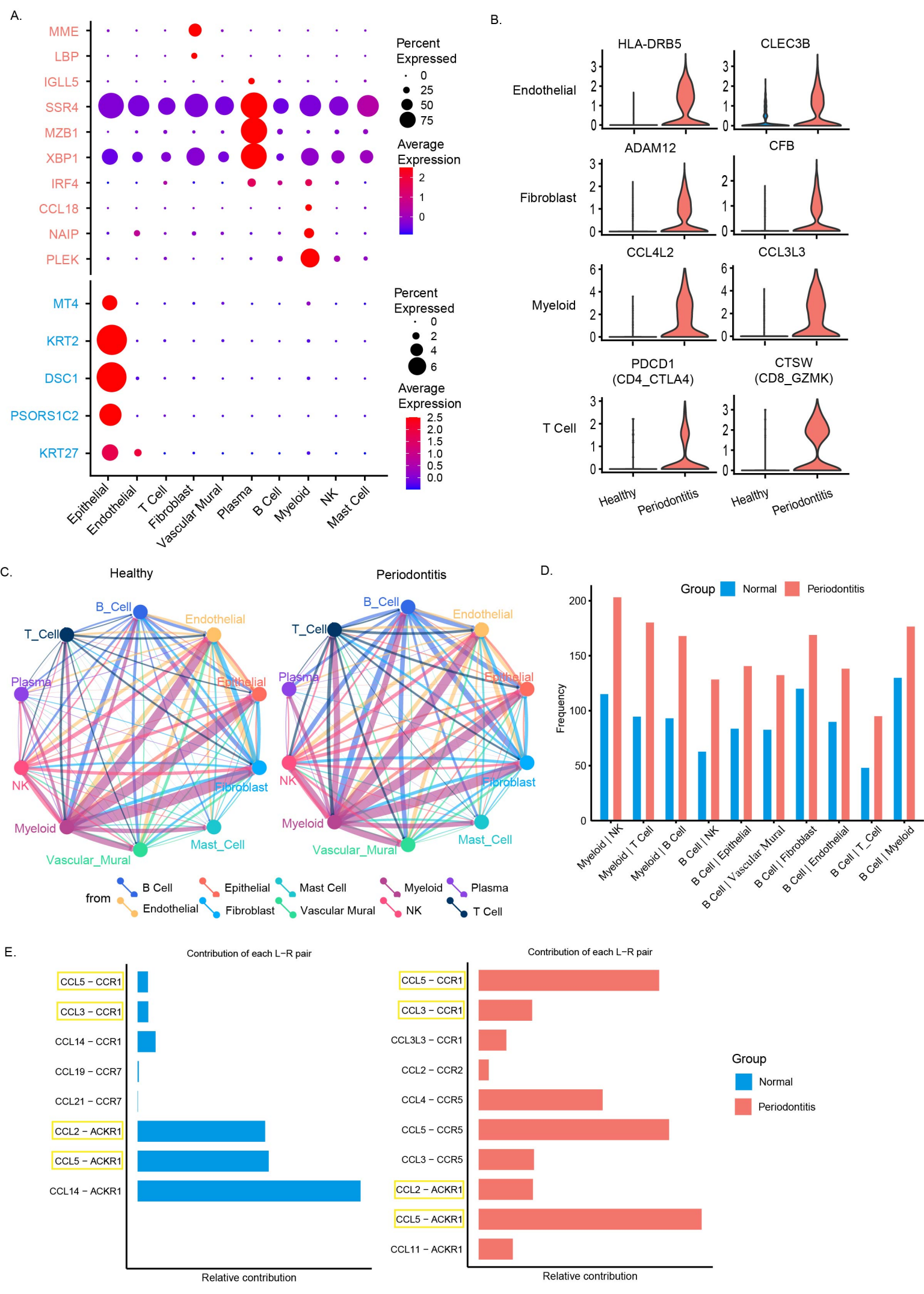

A.
Enamel
Gingival Epithelium
Oral Sulcular Epithelium
Junctional Epithelium
Dentin
Severe Inflammation
Connective Tissue
Biofilm
Tooth
Alveolar Bone
Healthy
Periodontitis
T cell
NK cell
Fibroblast cell
Plasma cell
B cell
Mast cell
Myeloid cell
Microorganism
B.
T Cell
NK
Endothelial
Epithelial
Vascular Mural
B Cell
Myeloid
Plasma
Mast Cell
Fibroblast
UMAP_1
UMAP_2
Nor1
Nor2
P1
P2
C.
KRT14, CDH1, EPCAM, KRT76, KRT16
PECAM1, CLDN5, VWF, CDH5, CD34
RGS5, PDGFRB, MCAM , ACTA2, TAGLN
DCN,COL1A1, COL1A1, COL3A1, LUM
CD79A, IGLC2, IGHG1, IGHG2, IGHG3
MS4A1, CD74, CD37,CD19
MS4A2, KIT, TPSB2, CPA3, TPSAB1
LYZ,CXCL8,HLA-DRA
CD3D, CD4, CD8A, IL7R,CD2
NKG7, GNLY
Epithelial
Endothelial
Vascular Mural
Fibroblast
Plasma
B Cell
Mast Cell
Myeloid
T Cell
NK
D.
Freq
100%
75%
50%
25%
0%
Nor1
Nor2
P1
P2
Epithelial
Endothelial
Vascular Mural
Fibroblast
Plasma
B Cell
Mast Cell
Myeloid
T Cell
NK
E.
Percent Expressed
0
25
50
75
Average Expression
−1 0 1 2
DSP
C1S
C1R
ANTXR2
FAM20A
SOS1
REST
ITGB4
CTSC
Dilated cardiomyopathy with woolly hair, keratoderma, and tooth agenesis
Ehlers-Danlos syndrome, periodontal type
Ehlers-Danlos syndrome, periodontal type
Juvenile hyaline fibromatosis
Amelogenesis imperfecta / gingival fibromatosis syndrome
Fibromatosis, gingival
Fibromatosis, gingival
Epidermolysis bullosa, junctional, non-Herlitz type
Periodontitis 1, juvenile
Epithelial
Endothelial
Vascular Mural
Fibroblast
Myeloid
B Cell
Plasma
Mast Cell
T Cell
NK


A.
Healthy
Periodontitis
Junctional epithelium
Spinous
Basal
Granular
JE1
JE2
BAS1
BAS2
SPN1
SPN2
SPN3
GRN1
GRN2
GRN3
UMAP_1
UMAP_2
B.
ITGB4
KRT19
FDCSP
DST
KRT15
CXCL14
COL17A1
KRT5
KRT1
KRT10
DSG1
CDH1
IVL
TGM3
SLURP1
CALML5
C.
Healthy
Periodontitis
D.
Amelogenesis (ODAM,C4orf26)
Serum Amyloid A Family (SAA1,SAA2)
HLA-DRs (HLA-DRB1,HLA-DRA)
Mitosis (MKI67,UBE2C,TOP2A)
Detoxification (GSTA4,CYB5A)
Epidermal Differentiation Complex (SPRR3,,S100A7,CRNN)
Anti-inflammatory (SLPI,SPINK5, MUC21,SLURP1)
Anti-microbial (DEFB1,S100A8,S100A9)
E.
ZNF774
MEF2A
RUNX2
AHR
BHLHE22
TCF7L1
TP63
GLIS1
E2F7
SP9
E2F2
MEIS2
U2AF1
VEZF1
POU2F2
SNAI1
EN1
ALX3
ETV2
DLX3
ZFHX2
GRHL3
SP5
ZNF681
NFE2L3
PPARD
NKX6.2
FOXC1
Scaled z-score
F.
Pseudotime
Componet_1
Componet_2
G.
Expression
Cluster
JE
BAS
SPN
GRN
CXCL1
FDCSP
IFITM1
CYR61
ITGB4
KRT15
NASP
ID3
HNRNPH1
IGFBP2
ID2
PTN
KRT5
SLC2A1
BTG2
KRT1
DSG1
LY6D
LGALS7
CRABP2
FABP5
IVL
CALML5
KRT10
S100A8
DEFB1
SLURP1
TGM3
SPRR2F
CRNN


A.
UMAP_2
UMAP_1
Endo_1
Endo_2
Endo_3
B.
ACKR1
SELE
SELP
HEY1
NOTCH4
GJA4
CCL21
PROX1
MMRN1
Expression Level
Endo_1
Endo_2
Endo_3
C.
Response to interferon gamma
Antigen processing and presentation
Leukocyte tethering or rolling
Regulation of blood pressure
Notch signaling pathway
Endothelium development
Endothelial cell migration
Regulation of vasculature development
Cell substrate adhesion
Lymph vessel development
ERK1 and ERK2 cascade
Lymphangiogenesis
GeneNumber
pvalue
D.
MHC Class II
(HLA−DRA,HLA−DRB1,HLA−DPA1,HLA−DPB1,HLA−DRB5)
Healthy
Periodontitis
UMAP_2
UMAP_1
E.
HLA-DR
CD31
Nuclei
Merge
50μm


A.
UMAP_2
UMAP_1
Fibro_1
Fibro_2
B.
CXCL13
IL32
SFRP2
OGN
RUNX2
PRELP
Expression Level
Fibro_1
Fibro_2
C.
Fibro_1 vs Fibro_2
Regulation of immune response
Regulation of cell adhesion
Defense response to other organism
Blood vessel morphogenesis
Defense response
Response to oxidative stress
Cell activation involved in immune response
Response to tumor necrosis factor
Response to interferon gamma
Regulation of MAPK cascade
Chondrocyte development
Osteoblast development
Bone remodeling
Ossification involved in bone remodeling
Detection of mechanical stimulus
Gap junction assembly
Amino acid transport
Positive regulation of morphogenesis of an epithelium
Negative chemotaxis
Keratan sulfate catabolic process
t value of GSVA scores
D.
Decorin
Osteoglycin
Nuclei
Merge
50μm
E.
Pearson Correlation
Epithelial.cell_KRT13.high.Adult.Esophagus2.
Keratioincyte.Adult.Esophagus1.
Mucosal.aquamous.epithelial.cell.Adult.Esophagus1.
Epithelial.cell_KRT4.high.Adult.Esophagus1.
Epithelial.cell_KRT16.high.Adult.Esophagus1.
Stromal.cell_PTGDS.high.Adult.Esophagus2.
Fibroblast.Adult.Esophagus2.
Fibroblast_.CXCL12.high.Adult.Artery1.
Myofibroblast_POSTN.high.Adult.Artery1.
Fibroblast.Adult.Lung2.
PV1.Placenta_VentoTormo.
Smooth.muscle.cell.Adult.Artery1.
Smooth.muscle.cell_CCL19.high.Adult.Trachea2.
Quiescent.stellate.cell.Adult.Pancreas_Baron.
Smooth.muscle.cell.Adult.Lung1.
Endothelial.cell.Testis_Guo.
Endothelial.cell_VWF.high.Adult.Artery1.
Endothelial.cell_VWF.high.Adult.Lung2.
Endo..m..Placenta_VentoTormo.
Endothelial.cell_SELE.high.Adult.Lung2.
Epithelial
Endothelial
Fibroblast
Mural Cell


A.
pDC_LILRA4
cDC1_BATF3
cDC2_CD1C
Mono_CD14
Macro_NLRP3
Macro_PRDM1
Macro_C1QA
UMAP_2
UMAP_1
B.
100%
75%
50%
25%
0%
Healthy
Periodontitis
pDC_LILRA4
cDC1_BATF3
cDC2_CD1C
Mono_CD14
Macro_C1QA
Macro_PRDM1
Macro_NLRP3
C.
P N P N P N P N P N P N P N
LILRA4
GZMB
JCHAIN
BATF3
CLEC9A
CADM1
CD1C
CLEC10A
FCER1A
CD14
C1QA
SEPP1
APOE
FCGR2B
PRDM1
HES1
NLRP3
IL1B
S100A8
pDC_LILRA4
cDC1_BATF3
cDC2_CD1C
Mono_CD14
Macro_C1QA
Macro_PRDM1
Macro_NLRP3
Percent Expressed
0
25
50
75
100
Average Expression
2
1
0
−1
Group
Normal
Periodontitis
D.
Phagocytosis
(MRC1,CD163,MERTK,C1QB)
Angiogenesis
(VEGFA,VCAN,CXCL8,ANGPTL4)
2.0
1.5
1.0
0.5
0.0
Macro_C1QA
Macro_PRDM1
Macro_NLRP3
Macro_C1QA
Macro_PRDM1
Macro_NLRP3
Macro_C1QA
Macro_PRDM1
Macro_NLRP3
E.
Velocity_UMAP
Macro_NLRP3
Macro_PRDM1
Macro_C1QA
F.
CD4_CTLA4
CD4_CCR7
CD4_FOS
CD8_GZMK
NK_GNLY
UMAP_2
UMAP_1
G.
100%
75%
50%
25%
0%
Healthy
Periodontitis
CD4_CTLA4
CD4_CCR7
CD4_FOS
CD8_GZMK
NK_GNLY
H.
P N P N P N P N P N P N
TNFRSF18
TNFRSF4
CTLA4
BATF
CCL20
CD28
ICOS
CCR7
GPR183
KLRB1
IL7R
FOS
JUN
PPP1R15A
DUSP1
GZMK
CCL5
CCL4
CCL4L2
CCL3L3
GNLY
TYROBP
KLRC1
XCL1
XCL2
CD4_CTLA4
CD4_CCR7
CD4_FOS
CD8_GZMK
NK_GNLY
Expression
2
1
0
−1
Cluster
CD4_CTLA4
CD4_CCR7
CD4_FOS
CD8_GZMK
NK_GNLY
Group
Normal
Periodontitis


A.
MME
LBP
IGLL5
SSR4
MZB1
XBP1
IRF4
CCL18
NAIP
PLEK
MT4
KRT2
DSC1
PSORS1C2
KRT27
Epithelial
Endothelial
T Cell
Fibroblast
Vascular Mural
Plasma
B Cell
Myeloid
NK
Mast Cell
Percent Expressed
Average Expression
B.
HLA-DRB5
CLEC3B
ADAM12
CFB
CCL4L2
CCL3L3
PDCD1 (CD4_CTLA4)
CTSW (CD8_GZMK)
Endothelial
Fibroblast
Myeloid
T Cell
Healthy
Periodontitis
C.
Healthy
Periodontitis
B_Cell
T_Cell
Endothelial
Plasma
Epithelial
NK
Fibroblast
Myeloid
Mast_Cell
Vascular_Mural
from
B Cell
Endothelial
Epithelial
Fibroblast
Mast Cell
Vascular Mural
Myeloid
NK
Plasma
T Cell
D.
Group
Normal
Periodontitis
Frequency
Myeloid | NK
Myeloid | T Cell
Myeloid | B Cell
B Cell | NK
B Cell | Epithelial
B Cell | Vascular Mural
B Cell | Fibroblast
B Cell | Endothelial
B Cell | T_Cell
B Cell | Myeloid
E.
Contribution of each L−R pair
CCL5 − CCR1
CCL3 − CCR1
CCL14 − CCR1
CCL19 − CCR7
CCL21 − CCR7
CCL2 − ACKR1
CCL5 − ACKR1
CCL14 − ACKR1
Relative contribution
CCL5 − CCR1
CCL3 − CCR1
CCL3L3 − CCR1
CCL2 − CCR2
CCL4 − CCR5
CCL5 − CCR5
CCL3 − CCR5
CCL2 − ACKR1
CCL5 − ACKR1
CCL11 − ACKR1
Relative contribution
Group
Normal
Periodontitis

Figures

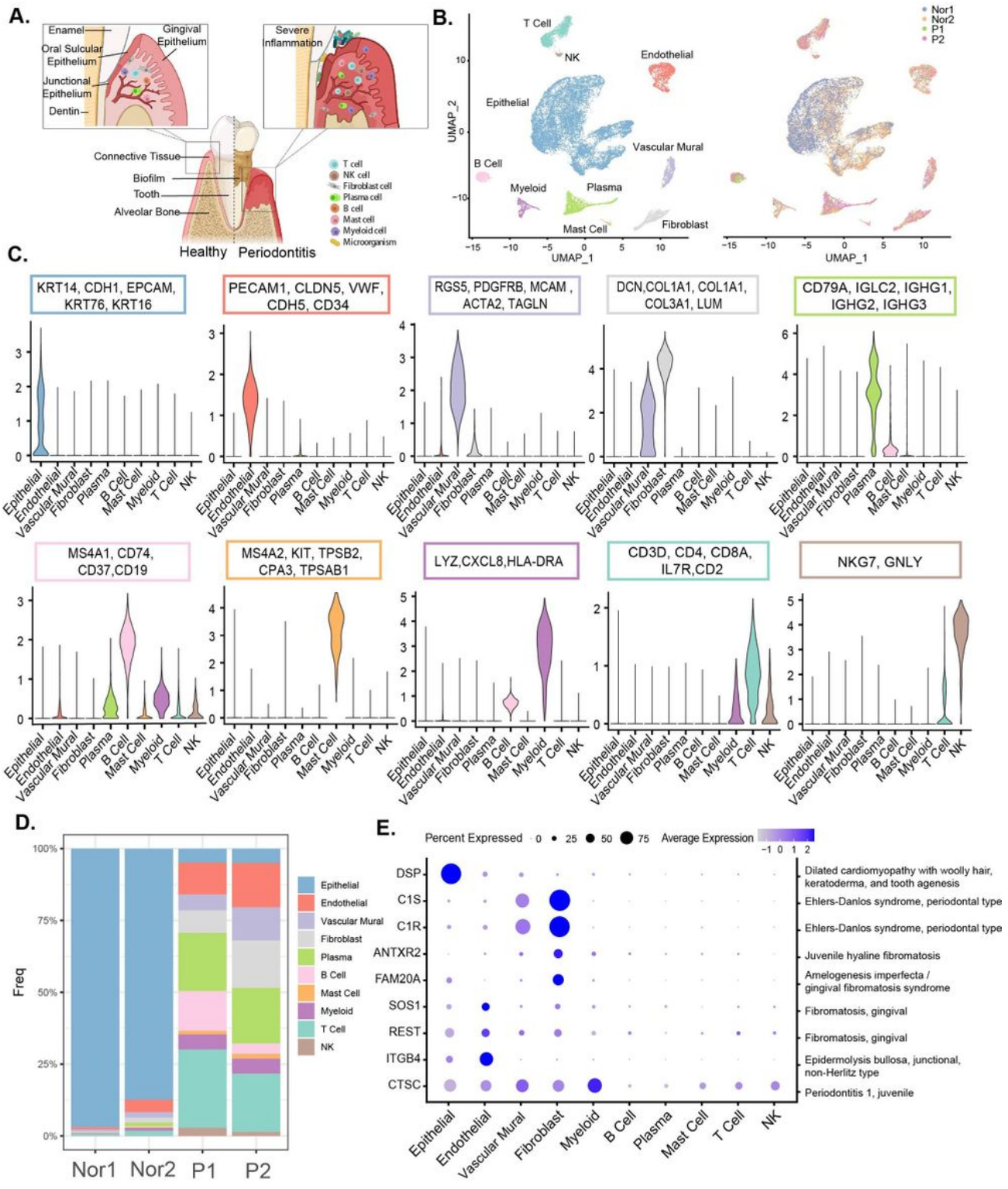


Figure 1

Overview of the clustering and annotation of the single-cell RNA sequencing data for gingival tissues. A. Schematic depicting anatomical regions analyzed in this study (The graph was created with BioRender.com). B. tSNE of the 29,967 cells profiled here, with each cell color-coded for (left to right): the

corresponding patient and the associated cell type. C. Well-known markers scores for annotation of gingival cell types visualized by violin plot. D. Proportion of cell types from periodontitis and healthy conditions. (Nor1 and Nor2 are healthy donors; P1 and P2 are periodontitis patients) E. Dot plot depicting gene expression levels and percentage of cells expressing genes associated with periodontal disease according to the OMIM database.

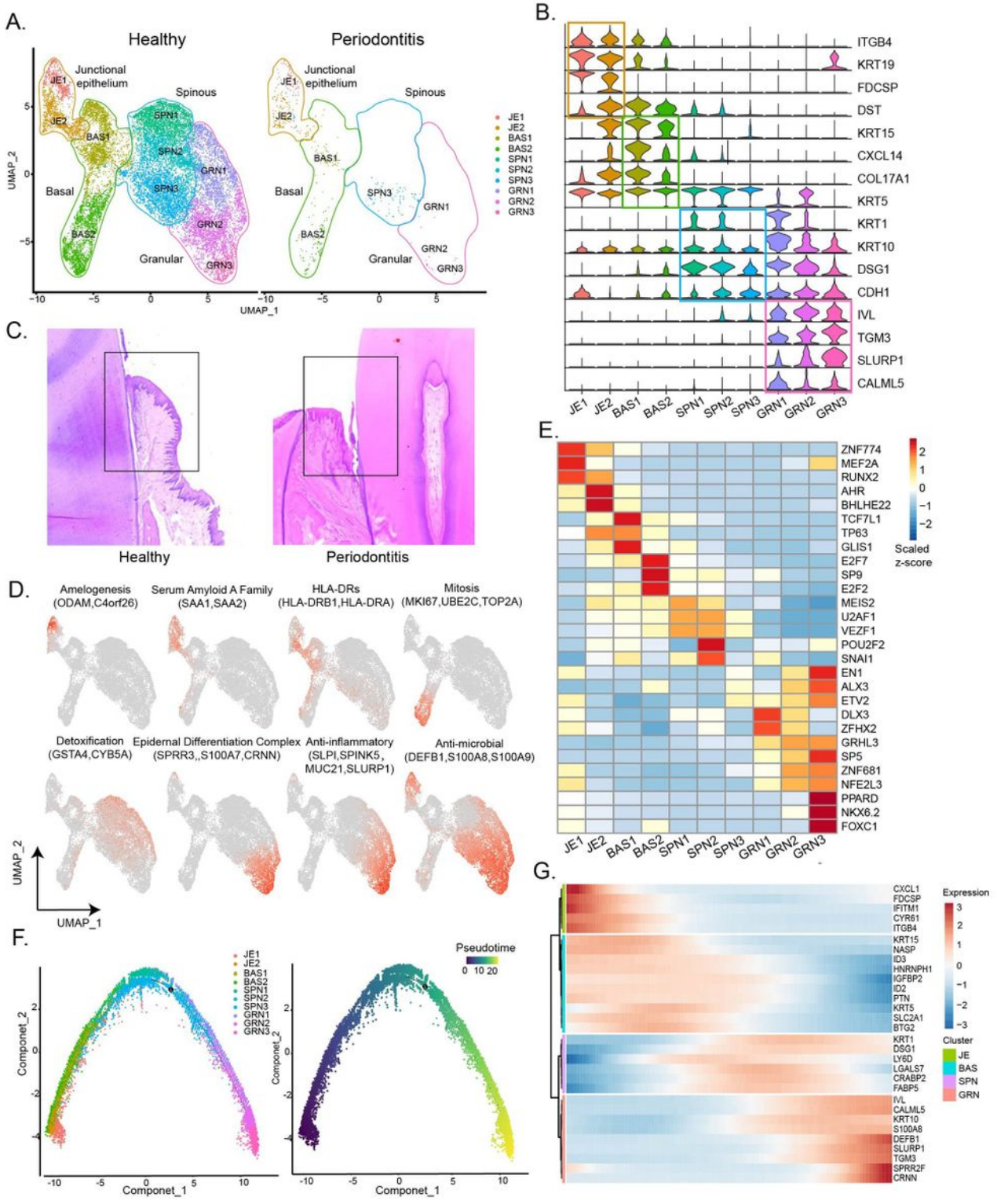


Figure 2

The heterogeneity and developmental trajectory of gingival epithelial. A. Visualization of human gingival epithelial cells in disease and healthy conditions separately. B. Markers for annotation of gingival epithelial cell types visualized by violin plot. C. Immunohistochemical showed the outermost layer of the epithelium experienced severe ruptures in patients. D. UMAP plots of epithelial cells, color-coded for functional molecules scores. E. Heatmap showing the activity of TFs in each epithelial cell subtype. The TF activity is scored using scaled AUCell. F. Monocle2-generated pseudotime trajectory of epithelial types imported from Seurat data, colored by cell-name designation (left) and colored in a gradient from dark to light yellow (right). G. Heatmap for clustering the 30 genes that vary as a function of pseudotime. The 30 genes were divided into four clusters, representing the genes at the beginning stage, the transitory stage, and the end stage of developmental trajectory, respectively.

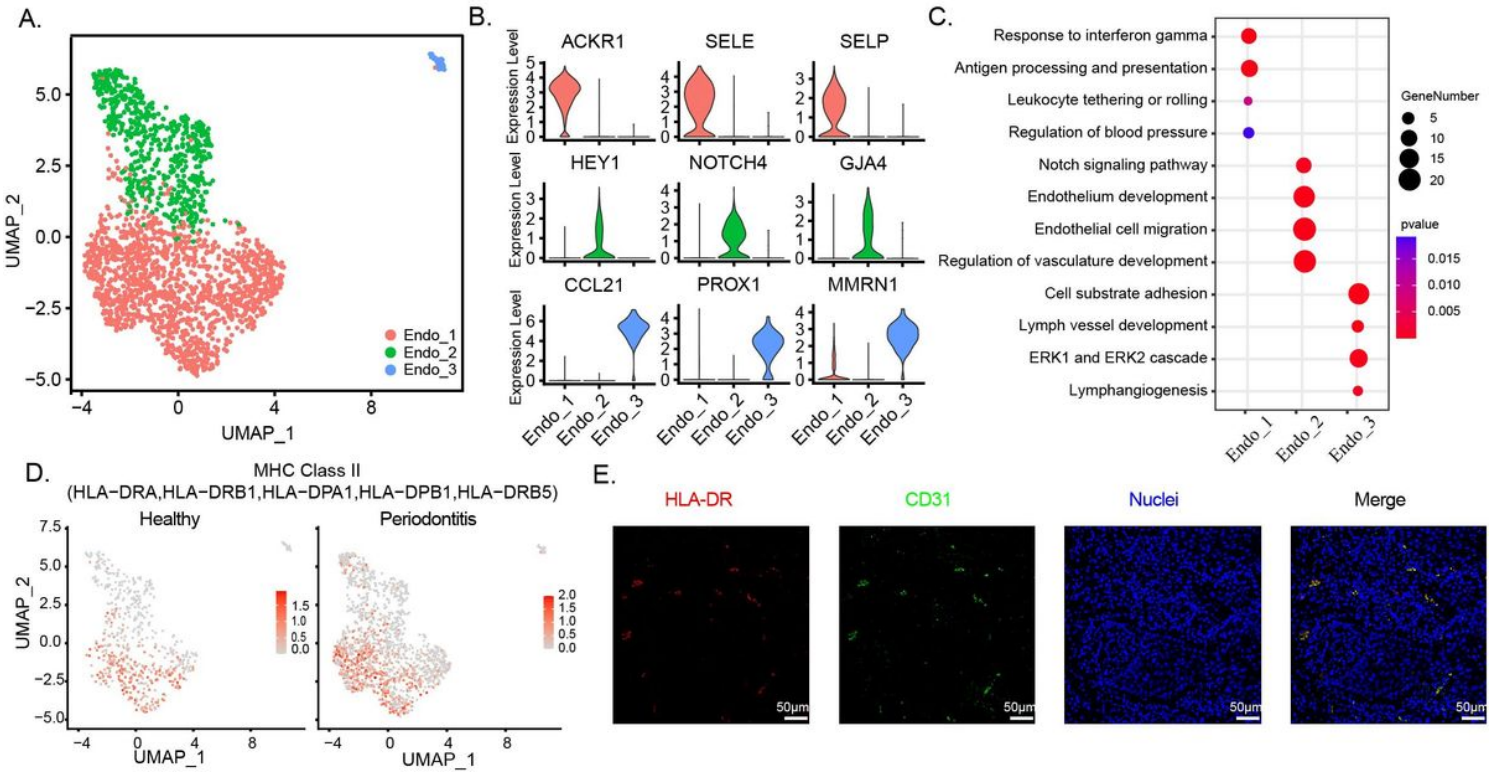


Figure 3

Three endothelial subclusters identified from the human gingival cells. A. UMAP visualization of three endothelial subclusters from human gingival cells. B. Violin plots showing the expression distribution of selected genes associated with functions in the endothelial clusters. C. Top gene ontology term enrichment for endothelial clusters. D. HLA-DRA showed expression specificity in the endothelial Endo_1. The gene expression levels are normalized and transformed as ln (CPM/10). E. Immunofluorescent (IF) staining validation of endothelial subtypes. Red color showed the signal of HLA-DR staining; green color showed the signal of CD-31 staining (endothelial marker); and blue color showed DAPI staining.

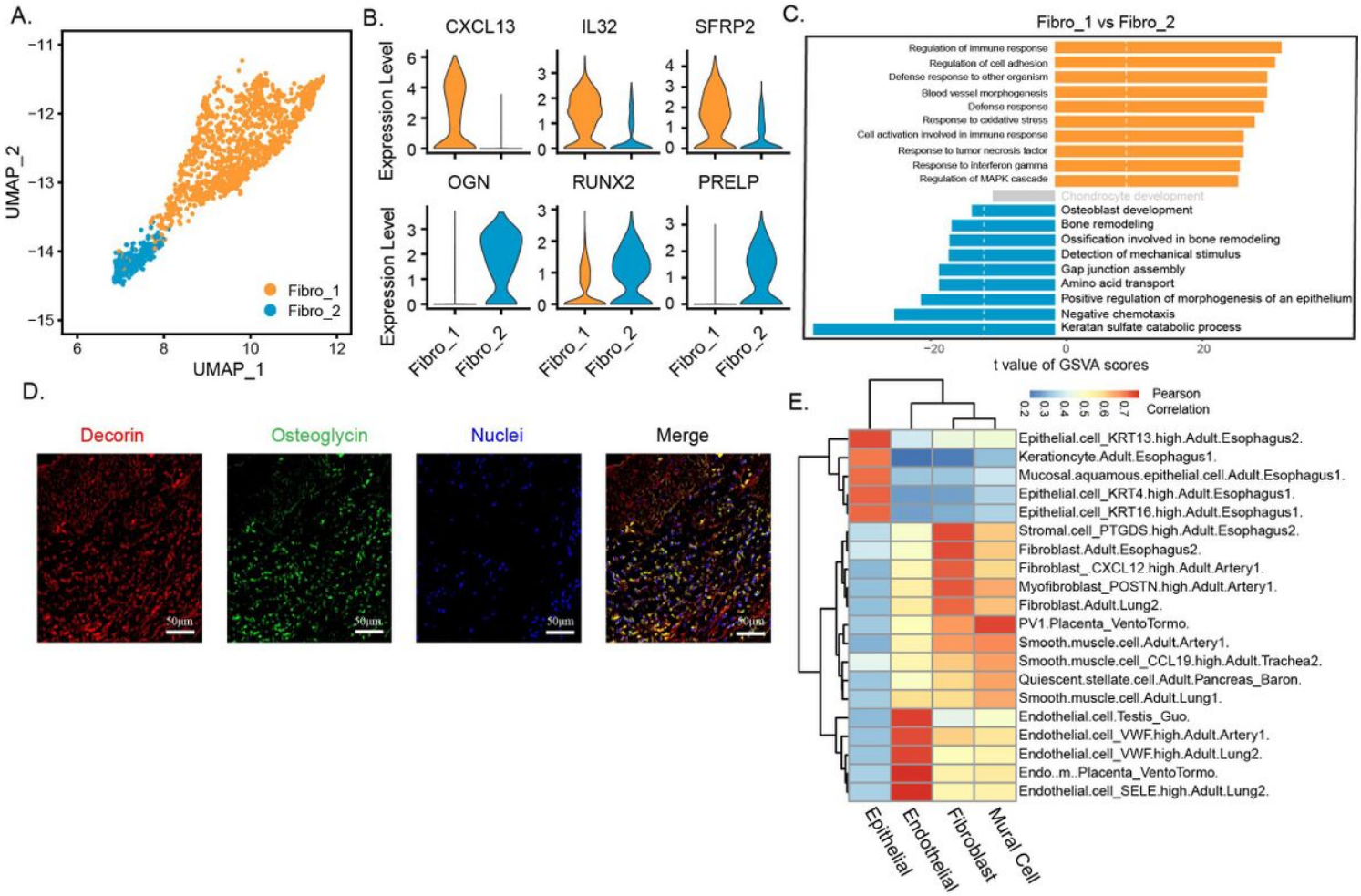


Figure 4

Two fibroblast subclusters identified from the human gingival cells. A. UMAP visualization of two fibroblast subclusters from human gingival cells. B. Violin plots showing the expression distribution of selected genes associated with functions in the endothelial clusters. The gene expression levels are normalized and transformed as ln (CPM/10). C. Differences in pathway activities scored per cell by GSVA between fibro_1 and fibro_2. Shown are t values from a linear model, corrected for fibro_1. D. Immunofluorescent (IF) staining validation of fibroblast subtypes. The IF was performed on subepithelial region and connective tissue, respectively. Red color showed the signal of 727 Decorin staining (fibroblast marker); green color showed the signal of Osteoglycin (OGN) staining; and blue color showed DAPI staining. E. Application of scHCL analysis for non-immune cells. Each row represents one cell type in scHCL. Each column represents a cell cluster in our dataset. Pearson correlation coefficient was used to evaluate cell-type gene expression similarity. Red indicates a high correlation; blue indicates a low correlation.

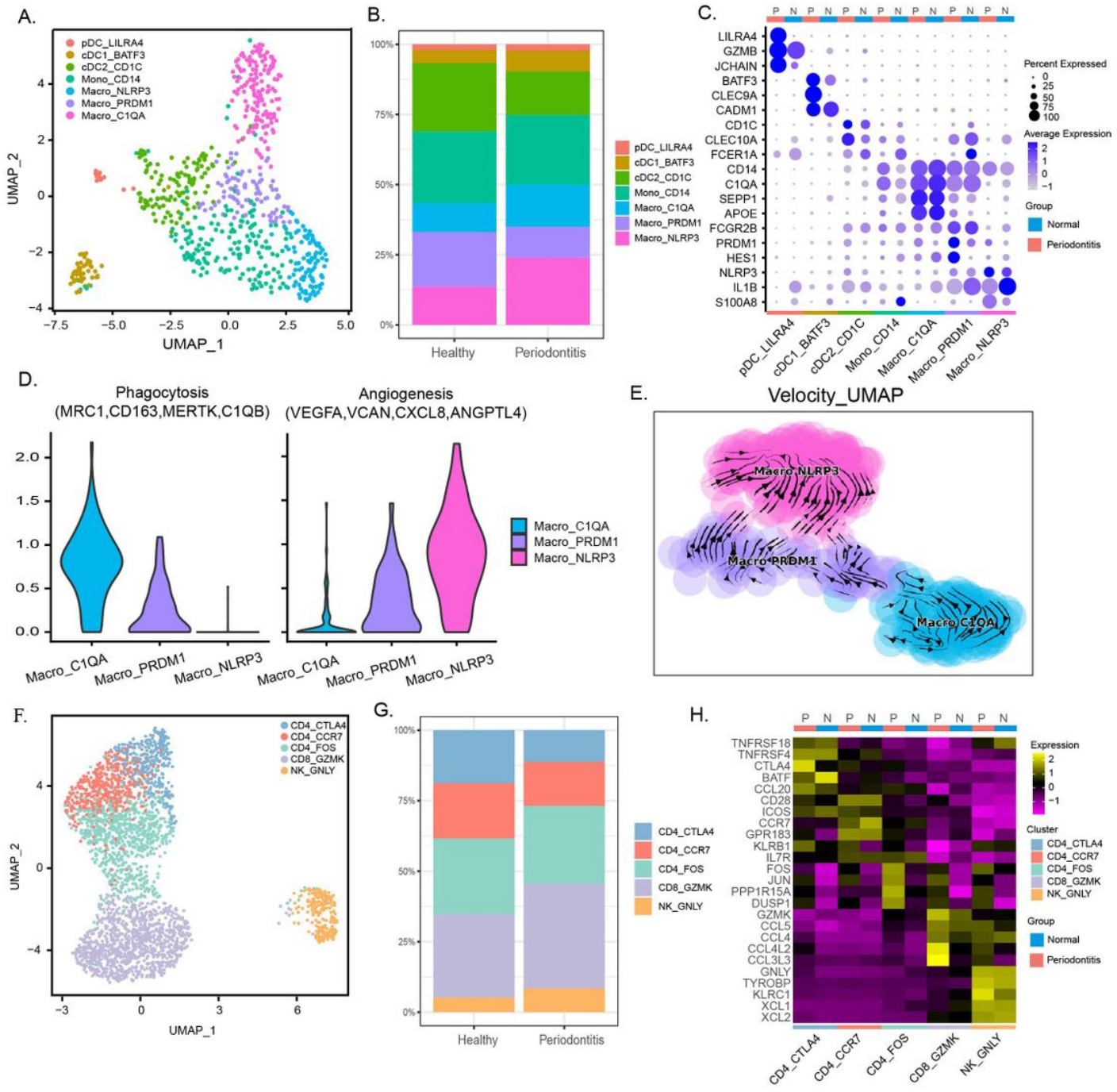


Figure 5

The heterogeneity of myeloid and T cells. A. UMAP visualization of seven myeloid clusters from human gingival cells. B. Relative proportion of myeloid cell subsets from different origin (healthy=2 and periodontitis=2). C. Bubble heatmap showing marker genes across seven myeloid clusters from (A). Dot size indicates fraction of expressing cells, colored according to z-score normalized expression levels. D. Violin plots showing the expression of angiogenesis- and phagocytosis-related genes in three macrophages clusters. The gene expression levels are normalized and transformed as ln (CPM/10). E. Velocity field projected onto the UMAP plot of three macrophages (arrows represent average velocity). F. UMAP visualization of seven myeloid clusters from human gingival cells. G. Relative proportion of T cell

subsets from different origin (healthy=2 and periodontitis=2). H. Heatmap of T cell top 5 DEGs in each cell cluster in healthy and periodontitis.

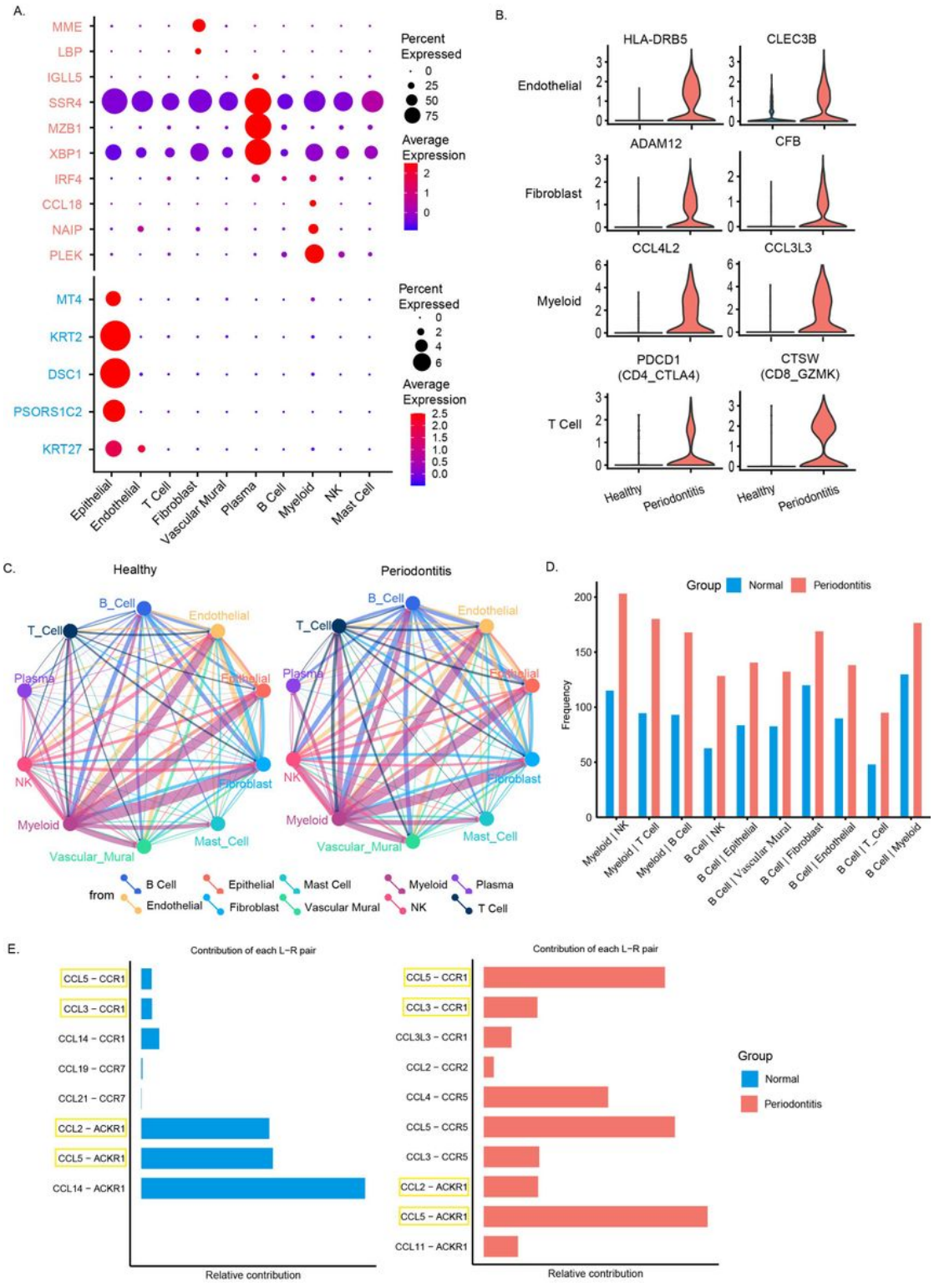


Figure 6

The difference among gingival cell types in healthy and periodontitis conditions. A. Bubble heatmap showing reported genes associated diseases clusters from (Figure 1B). Rose stands for reported increase in periodontitis, and turquoise stands for reported decrease in periodontitis. Dot size indicates fraction of

expressing cells, colored according to z-score normalized expression levels. B. Violin plots showing the expression of gene specific increasing in different clusters of periodontitis. The gene expression levels are normalized and transformed as ln (CPM/10). C. Cell-cell interaction network in healthy and periodontitis condition. Colors and widths of edges represent number of interaction pairs between cell types. D. The top 10 cell-cell interactions with the most significant increase in patients than healthy. The relative contribution of receptor-ligand in CCL pathway in normal and periodontitis. Yellow indicates that the common receptor-ligand is present in both the patient and the normal.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1S13.pdf
- supplementtable2OMIM.xlsx
- supplementtable3scenicTF.xlsx
- Supplementtable4monoclegenes.csv