Spatiotemporal cell landscape of human embryonic tooth development

Abstract

Introduction: Understanding the cellular composition and trajectory of human tooth development is valuable for dentistry and stem cell engineering research. Previous single-cell studies have focused on mature human tooth and developing mice tooth, but cell landscape on human embryonic dental development is still lacking.

Objective: We aimed to construct the spatiotemporal cell atlas of aborted fetus tooth germ.

Methods: We collected tooth germ tissues from aborted fetus (17-24 week) for single cell RNA sequence and spatial transcriptome. Subsequent clustering, spatial projection, pseudotime, gene regulation network, pathway enrichment and signaling network analysis were applied to reveal the cellular composition as well as its biological significance.

Results: We classified all cells into seven subclusters of epithelium, seven clusters of mesenchyme and other cell types like Schwann cell precursor and pericyte. For epithelium, the matrix cell-striatum intermedium branch and the ameloblast branch diverged from a same set of KRT15+-HOPX+-ALCAM+ epithelial stem cell lineage, but the spatial distribution of two branches were not clearly distinct. This trajectory received spatially adjacent regulation signals from mesenchyme and pericyte, including JAG1 and APP. The differentiation of pulp cell and pre-odontoblast showed four waves of temporally distinct gene expression, which involved regulation networks of LHX9, DLX5 and SP7 and were regulated by upstream ligands like BMP family.

Conclusion: We provided a reference landscape for the research on human early tooth development, covering different spatial structures and developmental periods.

Introduction

Human tooth development is a long and complex process starts at embryonic stage and last until adolescent period. This process represents a good model of organogenesis involving epithelial-mesenchyme interaction, morphogenesis and adult tissue self-renewal[1]. At the embryonic stage, dental development started at the formation of mesenchyme from neural crest[2] and the dental placode originated from the thickened oral epithelium. At this stage, dental epithelium starts to proliferate in some regions of the dental lamina, which then migrates to the deep connective tissue, and the end-stage epithelium proliferate and further differentiate to form the enamel organ. In parallel to this process, the ectomesenchyme layers beneath the proliferative epithelium also proliferate rapidly and aggregate around the epithelial layer to form the condensed mesenchyme layer. The epithelia and mesenchyme that proliferate in this restricted regions together consist in the tooth germ[3], which is the structural basis of further dental development The tooth germ is made up of three parts: 1. Oral ectoderm-originated enamel organ, which give rise to enamel; 2. Ectomesenchyme-originated dental papilla tissue, which give rise to dentin and pulp

tissue; 3. Ectomesenchyme-originated dental follicle, which give rise to cementum, periodontal ligament and alveolar bone[4]. A comprehensive understanding of dental development would require a cell atlas that cover all these heterogenous cell types and tissues.

The rapid development of single cell techniques provides new opportunity for this task[5]. Recent single-cell studies on mouse dental tissues have elucidated the cell heterogeneity from early state Cranial neural crest cells[6] to later incisor tissues[7], and delineated the differentiation and migration trajectory of epithelial and mesenchymal cells. Other studies using dental tissues from healthy volunteers also constructed dell landscape of postnatal human dental pulp[8,9], apical papilla[7], periodontal tissues[8,10] and immature tooth germ[11]. Based on these and other existing data, Hermans et al.[12] built a comprehensive dental cell landscape that basically covered all developmental stages of mouse incisor and molar tissue as well as human postnatal dental tissues, which has been proved valuable for clinical and bioengineering studies. However, some challenges remain unsolved by these existing efforts. An obvious blank is the lack of data on human prenatal dental tissues. Since mouse incisor has everlasting self-renewal capacity and exhibits dramatic physiological difference compared with human[13], the immediate translation of dental development knowledge from mouse to human should be interpreted with caution. Another drawback is the insufficient spatial resolution: current single-cell analysis relied on immunofluorescence that tags one or two proteins to study the spatial distribution of cell subtypes, which inevitably missed more refined information on developmental process based on multiple genes.

To fill in these blanks, we collected primary and permanent tooth germ tissues from 17 to 24 gestational week aborted fetus and applied single-cell RNA sequencing (scRNA) and spatial transcriptome (ST) to construct a spatiotemporal cell landscape of early human tooth development. Our study aimed to address the following questions. First, the cellular composition of human embryonic tooth germ at different stages. Second, the developmental trajectory of dental epithelium and mesenchyme, as well as their underlying biological process, gene regulatory networks and intercellular signaling pathways. Third, the spatial distribution of these elements.

Method

Sample collection and preprocessing

This study is approved and supervised by the ethic committee of Shanghai Tongren Hospital. Five fresh human fetal tooth germ tissues were collected from elective, nonmedically motivated pregnancy terminations at different embryonic developmental stages (17, 20 and 24 PCW), without evidence of developmental abnormalities. Informed consent was obtained from each donor before termination of pregnancy. Details of sample preprocess was provided in Supplementary methods.

scRNA and ST

Library construction, sequencing and data filtration sere described in the Supplementary methods. Both libraries of scRNA and ST were sequenced with paired-end 150 bp sequencing (PE150) by NovaSeq 6000 platform. Library size normalization was performed with NormalizeData function in Seurat[14] to obtain the normalized count. Specifically, the global-scaling normalization method “LogNormalize” normalized the gene expression measurements for each cell by the total expression, multiplied by a scaling factor (10,000 by

default), and transformed by Seurat: SCTransform function. Cell phase was defined by Seurat: CellCycleScoring function based on cell cycle gene list[15]. Spatial transcriptome data were preprocessed by Seurat in a similar way.

Cluster analysis

For scRNA data, we applied Louvain community detection-based cluster analysis implemented in Seurat: FindClusters function with number of principle component=20 and resolution=0.5. We applied Seurat: FindMarkers Wilcoxon test to find top 15 markers for each cluster, and defined them according to existing knowledge on dental cell types. Epithelium and mesenchyme cell were further extracted and underwent second round of clustering, with the same parameters.

Pseudotime analysis

For epithelium and mesenchyme (excluding isolated follicle and MSC clusters) cell groups, we applied monocle3:learn_graph and order_cells function[16] to learn the pseudotime trajectory. We then applied monocle3: graph_test function on top 5,000 variable genes detected by Seurat: FindVariableGenes function to find all genes that significantly altered along the trajectory. Genes with FDR-adjusted $p<0.01$ were further used for module detection by monocle3: find_gene_modules function. For the ease of interpretation, in epithelium cell group we separately applied module detection in each branch, and only retained gene modules that showed up-regulation or down-regulation along each of the three branches.

Transcription factor regulation network analysis

We applied SCENIC[17] to construct transcription factors (TF)-target genes regulation network, separately for epithelium and mesenchyme. TF-motif annotation was downloaded from SCENIC website, which was corresponded to hg38 refseq gene annotation and contained a 10kb window of each gene. We retained target genes for extended network of each TF, and applied Fisher test to see if module genes detected by pseudotime analysis were enriched in any of the TF regulation network. TFs with FDR-adjusted $p<0.01$ and Odds ratio >2 were retained for each module. We further extracted the overall expression of a TF and all its targets (so-called regulon) of each cell to verify that this expression value had the similar expression pattern as module genes.

Gene Ontology analysis

We used ClusterProfiler R package[18] to apply Gene Ontology-Biological Process (GO-BP) analysis by a hypergeometry test. The background gene list was set as all genes with GO-BP annotations. Only terms with >5 and <500 genes were considered. We applied simplify function to remove highly similar terms. Only terms with FDR-adjusted $p<0.01$ were retained. The odds ratio and p value were reported in the word cloud plot.

nichenetr analysis

We used nichenetr R package[19] to apply signaling network analysis as described previously[11]. We separately considered 1) up-regulation modules of each of the three epithelium branches; 2) immature (first and second wave) and mature (fourth wave) of mesenchyme differentiation as target gene list, and the final signaling network were combined in each of these two categories. Ligands were prioritized by area under curve (AUC) and Pearson correlation coefficient calculated by nichenetr.

Spatial transcriptome analysis

After defining cell clusters on scRNA data, we projected it to spatial transcriptome data by Seurat: FindTransferAnchors and TransferData functions to estimate the cell type probability of each spot. We assign cell types with the highest probability to each spot. We used misty R package[20] to estimate the spatial relationship among highlighted cell types, TFs and expression levels of ligands and target genes. We chose the "intra" mode in misty analysis and trained a linear model to predict each of the target values (cell probability or gene expression) by adjacent predicting values. Spatial adjacency was quantified by importance score in the prediction model.

Result

Overview of cell types in human embryonic tooth

We collected tooth germ tissues from aborted fetus spanning 17 to 24 gestational weeks (Figure 1A) for scRNA and ST. In scRNA, we obtained expression data for 11,218 single cells of nine general cell types after quality control (Figure 1B-D). As shown in Figure 1D and Figure S1, we classified these cell groups based on existing cell atlas of mouse tooth[7], adolescent tooth[8] and histology knowledges. We first defined cells with epithelial origin based on expression of epithelial stem cell marker like KRT15 and COL17A1, mature epithelium markers like EDN2 and ameloblast markers like ASCL5 and AMELX. Mesenchyme origin cells were characterized by their expression of SCUBE1, NES and MSX1. An isolated subgroup of mesenchyme origin cells was defined as follicle cells based on expression of SFRP2 and GDF10. Lastly, we defined mitotic cells, pericyte, endothelium, Schwann cell precursor (SCP), macrophage and blood stem cells based on their specific markers CENPF, RGS5, CD34, FOXD3, C1QA and PTPRC, respectively (Table S1). Across different gestational weeks, we observed difference in the proportion of epithelium and pericytes, but no difference in cell cycle distribution (Figure 1C).

We further applied ST on tooth slices from three donors: a slice from 24-week embryo that contained both tooth germ of primary tooth and tooth germ of permanent tooth, a slice of 17-week primary tooth and a slice of 20-week primary tooth (Figure 1E). We applied anchor-based integration method implemented in Seurat v3[14] to evaluate the proportion of each cell cluster of scRNA in each spatial plot (Figure 1F and Figure S2). As expected, we observed that epithelium incompletely surrounded pulp tissues that were consisted mainly of mesenchyme cells and scattered pericytes and endothelium. Follicles were only observed in permanent tooth germ, in line with their slower maturation process compared with primary tooth. We also found an isolated SCP group surrounded by endothelium and follicle cells, which were adjacent to permanent tooth germ and likely represented the process of neuron outgrowth. The spatial distribution of cell markers of corresponding cell types also supported these spatial characteristics (Figure 1G). Taken together, we successfully constructed the cell type atlas of human embryonic tooth germ that covered various developmental timepoints and spatial structures.

Branched trajectory of dental epithelium differentiation

Next, we conducted in-depth analysis of the cell constitution and development process of epithelium cell groups. By applying cluster analysis and pseudotime analysis, we partitioned

all epithelial origin cells into three branches (Figure 2A). The root branch expressed COL17A1, and consisted of early stem cells expressing KRT15 (KRT15+ SC, Figure 2B) and late stem cells expressing IGFBP5, HOPX and FGL2 (HOPX+ SC). Root branch ended at ALCAM positive stem cells, which then gave rise to two branches: the SI branch consisted of matrix cells expressing KRT7 and TAGLN and striatum intermedium expressing CLDN10, and the ameloblast branch expressing ASCL5 and AMELX (Figure S3). Lastly, we found a small subcluster of ameloblast expressing WNT10A, which we defined as differentiated ameloblast (diff amelo) due to the known role of Wnt signals in suppressing ameloblast development[21]. 24-week tooth germ contained moderately more ameloblast and less ALCAM+ SC than 17-week tooth germ, and the overall cell proportion was similar (Figure 2C). These results suggested that the development of SI have diverge trajectory, even though they share a same set of epithelial stem cells as progenitor.

We then analyzed the biological process underlying these branches. At the threshold of FDR-adjusted $p<0.01$, we found 622, 1397 and 842 genes whose expression levels significantly altered along pseudotime in root, SI and ameloblast branch, respectively. We applied gene module detection algorithm implemented in monocle3[16] to find genes that were consistently up- or down-regulated in each branch (Figure 2D and Table S2-4). In root branch, gene related to translational initiation (Gene Ontology [GO] adjusted $p=4.31\times10^{-5}$) and epithelial cell proliferation (GO adjusted $p=8.95\times10^{-5}$) were significantly enriched in up-regulation gene module (Figure 2E), suggesting that these biological processes were activated. Similarly, we found activation of regulation of MAP kinase activity (GO adjusted $p=1.60\times10^{-4}$) and response to BMP (GO adjusted $p=3.26\times10^{-4}$) during striatum intermedium development, and activation of extracellular matrix organization (GO adjusted $p=3.56\times10^{-6}$) and connective tissue development (GO adjusted $p=2.14\times10^{-5}$) during ameloblast development (Figure 2E). On the contrary, process of epithelial cell proliferation was suppressed as SI and ameloblast differentiated from root branch (Figure 2E).

Lastly, we ran SCENIC[17] to find transcription factors (TF) that regulate these biological processes. As shown in Figure 2F and Table S5-7, TF like NR1D1, DLX3, RUNX2, SOX11 and SMAD9 were predicted to regulate a large number of genes showing alterations in each branch. For example, among 18 target genes of DLX3, 14 were down-regulated in the root branch (Odds ratio, OR=102, Fisher test $p=6.45\times10^{-18}$). Consistently, the overall expression levels of DLX3 and its targets (so-called regulon[17]) were highly expressed in the primitive KRT15+ SC, and gradually diminished in the root branch (Figure 2G). We highlighted regulons of several other TF showing consistent expression patterns in SI and ameloblast branch, such as SOX11 and SMAD9 that were activated during SI and ameloblast development (Figure 2G).

Spatial patterns of dental epithelium development

We next linked the single cell-level signals to the spatial transcriptome. As shown in Figure 3A, B and Figure S4, KRT15+ SC were only found in the outer surface of permanent tooth germ, whereas ALCAM+ SC were found inner to KRT15+ SC in permanent tooth germ.

Consistently, ameloblast were only sparsely distributed inner to stem cell layers in permanent tooth germ, but were abundant and densely distributed in primary tooth germ. Ameloblast were found at the interface between pulp tissue and epithelial tissues, whereas matrix cells were distributed at the outer layers of ameloblast. These patterns were also supported by the spatial expression levels of the corresponding cell markers (Figure S5). We also summarized the inferred cell proportion of each spatial spot of all tissues (Figure 3C), and found that epithelial stem cells were mainly distributed in permanent tooth germ, whereas striatum intermedium were mainly found in tooth germ of 24 gestational week.

We then applied misty[20] to elucidate the spatial adjacency of different epithelial cell subtypes and key TFs. As shown in Figure 3D, when analyzing epithelial cells along, KRT15+ SC and ALCAM+ SC had close distribution with each other, whereas ameloblast, differentiated ameloblast and striatum intermedium aggregated together. Including all cell types together, only mesenchyme showed close relationship with multiple epithelial cells. When analyzing key TFs in epithelial differentiation (i.e., those TFs highlighted in Figure 2), we observed similar pattern that only adjacent mesenchyme cell proportion had a profound impact on multiple TF expression levels (Figure 3D). In general, we inferred that signal from mesenchyme cells had an important role in regulating dental epithelia differentiation.

Involvement of LHX9, TCF7 and SP7 in dental pulp development

Similar to the analysis of dental epithelium, we further deciphered the spatiotemporal dynamics of mesenchyme cell group. As shown in Figure 4A-C and Figure S6, we identified seven subtypes of mesenchyme origin. Apical progenitors (AP) were defined based on the expression of SFRP1, SOSTDC1 and SMOC2, which were further divided according to the expression of PENK, a gene encoding precursor of enkephalin (PENK+ and PENK- AP, respectively). We defined pulp cells according to expression of TWIST2 and KCNIP4, and defined pre-odontoblast according to DLX3 and SALL1 (SALL1+ OD). Another subtype of pre-odontoblast was characterized by the expression of DKK1 and IGFN1 (IGFN1+ OD). Lastly, an isolated and likely quiescent group of progenitor cells highly expressed TNC and HEY1, which we defined as mesenchymal stem cell (TNC+ MSC) according to previous cell atlas[8]. Compared with 24-week tooth germ, 17-week tooth germ had significantly larger proportions of follicle cells and apical progenitor, and significantly smaller proportion of pulp cell, pre-odontoblast and MSC (Figure 4C).

We further applied pseudotime analysis to reconstruct the developmental trajectory from apical progenitor to pre-odontoblast (Figure S7), and identified four waves of gene expression patterns along this trajectory (Table S8). As shown in Figure 4D, the first and second wave corresponded to progenitor cell differentiation, and the third and fourth wave corresponded to pulp cell development. The first wave included HES1 and TWIST1, and were enriched in cell fate determination (GO p-adjust=2.28×10^{-3}) and other processes related to cell cycle (Figure 4E). SCENIC[17] analysis revealed that TFs regulating the first wave genes were immediate genes like FOS and JUN family like FOSB (Fisher test p=3.60×10^{-14}; Figure 4E and Table S9). The second wave included genes like PAX9, SCUBE3 and SPARC, and were enriched in extracellular matrix organization (GO p-adjust=1.23×10^{-18}) and formation of primary germ

layer (GO p-adjust=8.38×10^{-9}). They were also enriched in the predicted targets of LHX9, TCF7 and SOX18 (Fisher test $p<10^{-10}$; Figure 4E and Table S9). The third wave consisted of genes taking part in cytoplasmic translation (GO p-adjust=7.66×10^{-29}), without significant enrichment in specific TF regulon. Lastly, the fourth wave genes, including DKK2, VCAN and THY1, were involved in regulation of odontogenesis (GO p-adjust=1.68×10^{-4}), regulation of Wnt signaling pathway (GO p-adjust=0.0001) and ossification (GO p-adjust=3.18×10^{-6}). They were mainly regulated by TF like DLX5,DLX3, SP7 and MSX2 (Fisher test $p<1.78\times10^{-5}$; Figure 4E and Table S9). The overall expression levels of regulons of these highlighted TF were also in line with the four wave patterns (Figure 4F).

Pulp cell subtypes had distinct spatial distribution

We applied Seurat anchor-based integration algorithm[14] to map single cell data to spatial transcriptome (Figure 5A). In 17-week primary tooth germ, we observed a PENK+ progenitor →PENK- progenitor→pulp cell→epithelium layer distribution from distal to proximal end, in line with their order in the developmental trajectory (Figure 4A). In the 24-week primary tooth germ, we further discovered that SALL1+ pre-odontoblast was at the outer layer of pulp cell, whereas IGFN1+ pre-odontoblast was aggregately located within pulp cell cluster. One exception was the TNC+ MSC, which was sparsely distributed in permanent tooth germ without significant spatial structure (Figure 5A and Figure S8). We plotted the expression of its marker TNC and HEY1 and also found no matching patterns (Figure 5B). We inferred that TNC+ MSC was loosely distributed in all regions of pulp tissues. As a comparison, the density as well as marker expression of other mesenchyme cell subtypes (Figure 5B and Figure S8 & 9) were all aggregated to specific regions. The temporal distribution (Figure 5C) showed that TNC+ MSC, follicle cell and PENK+ progenitors were mainly found in immature permanent tooth germ ($p<10^{-10}$), whereas pre-odontoblast and pulp cells were mainly found in mature primary tooth germ ($p<10^{-10}$).

We also applied misty[20] to quantify the spatial adjacency among cell subtypes as well as the key TFs. Within mesenchyme cell groups, we found close relationship between follicle cell and PENK+ progenitor (misty adjacency score>2.5) as well as SALL1+ OD and pulp cell (adjacency score>2). Outside mesenchyme cell groups, only epithelium and mitotic cells had close relationship with pulp cells (adjacency score>2). For TFs, pulp cells had close association with nearby expression values of multiple TFs that had important roles in mesenchyme differentiation, except FOSB which took effect in early progenitors. Another spatial association was found between SALL1+ OD and MSX2 (adjacency score=2.02).

Signals from mesenchyme and pericyte controlled dental epithelium development

Having discovered the refined cell subtypes of embryonic tooth germ, we were able to study the intercellular signaling pathways that regulate dental development. We first applied nichenetr[19] to build a signaling network that regulate branch-specific key genes in epithelial development (181 key genes included in nichenetr; Figure 6A). Integrating the prioritized

upstream ligands of both root branch, SI branch and ameloblast branch, we found that signals from mesenchyme (including follicle cell) and pericyte regulated the largest number of epithelial key genes (n=142 and 92, respectively, Figure 6A and Table S10). Top credible ligands included BMP5 (highest expression in SALL1+ OD, nichenetr area under curve [AUC]=0.62), SFRP2 (highest expression in follicle cell, AUC=0.58), JAG1 (highest expression in pericyte, AUC=0.57), HMGB2 (highest expression in macrophage, AUC=0.56, Table S11), etc. Among them, JAG1 was predicted to regulate the largest number (n=91) of epithelial key genes, including GJA1, NOTCH1, etc. Two notable exceptions were GPNMB and SHH: they were the only prioritized ligands that showed highest expression in ameloblast (Figure S10) and had high potential to regulate SI and ameloblast branch development (AUC=0.59 and 0.53, respectively), suggesting that they might act through autocrine or paracrine. Despite their expression in epithelium, the highest expression was found in SCP (Figure 6A).

We further applied misty on spatial transcriptome data to test whether these prioritized ligand-target pairs were also spatially adjacent. At the threshold of top 10% adjacency score (Method), 12 out of 19 prioritized ligands could impact the adjacent expression level of at least one targets (Figure 6B). Top ligands like SHH, BMP7 and APP had close relationship with more than ten predicted targets, and these targets covered all three branches of epithelial development. SFRP2 were also predicted to impact adjacent expression of nine targets. One example is GJA1 (Figure 6C): spatial expression of SFRP2 and GJA1 was close but not overlapped, which suggested that SFRP2 was secreted to regulate GJA1 in nearby cells via endocrine instead of paracrine or autocrine. We also showed the spatial characteristics of other ligands in Figure S11. On the other hand, some targets like ID1, ID3 and CCND1 were predicted to be regulated by more than five ligands both by nichenetr and by spatial adjacency, implicating complex crosstalk among different signaling pathways. Taken together, we integrated single-cell and spatial information to prioritize signaling pathways of SHH, BMP, APP and SFRP2 in the regulation of dental epithelium development.

BMP signaling pathways regulated mesenchyme differentiation in tooth germ

Lastly, we analyzed the intercellular signaling networks that regulated dental mesenchyme development. For all 162 mesenchyme key genes (genes of first, second and fourth wave of Figure 4D & E), 135 were regulated by ligands from pericyte, and 133 were regulated by ligands from epithelium (Figure 7A). These included ligands like BMP5 (predicted number of target n=122, AUC=0.75), BMP7 (n=116, AUC=0.72), DSC3 (n=94, AUC=0.75) and COL4A1 (n=57, AUC=0.72, Table S12 &13), etc. We also found that different ligands tended to regulate different stages of mesenchyme development. For example, BMP5 and BMP7 regulated all first wave genes and more than 70% of the second wave genes that were included in nichenetr analysis, whereas all predicted targets of APP (mainly expressed in endothelium) and TGFB1 (mainly expressed in macrophage) belonged to the fourth wave (Figure 7A). In general, upstream signaling pathways of mesenchyme development was restricted to a few vital ligands and cell types, such as BMP signals from pericyte and epithelium.

In spatial transcriptome analysis, we found that COL4A1 and BMP7 had close relationship with a large number of predicted targets (Figure 7B). COL4A1 could impacted the adjacent expression of ten target genes of first wave and 17 target genes of second wave,

whereas BMP7 could impact three of first wave and 18 of second wave. Beside COL4A1 and BMP7, only a few ligand-target pairs had significant spatial adjacency, such as APP (adjacent with four targets) and TGFB1 (adjacent with two targets). The spatial expression patterns (Figure 7C and Figure S12) also supported these results. We inferred that mesenchyme development was mainly regulated by a few upstream ligands, such as BMP family, COL4A1, DSC3 and TGFB1.

Discussion

In the current study, we applied single-cell RNA sequencing and spatial transcriptome on human embryonic tooth germ to depict a spatiotemporal cell landscape of human tooth development. We revealed the cellular composition as well as their spatial distribution in human embryonic tooth germ spanning different developmental periods and tissues from both primary and permanent tooth germ. We also depicted the developmental trajectory, biological insights, essential transcription factors and signaling pathways of dental epithelium and mesenchyme development. In complement with the existing dental cell atlas of prenatal and postnatal mouse tooth[6,7], postnatal human tooth[7,8] and tooth germ[11], our result brought one more step closer to the intact single-cell picture of dental development[12].

Our collection of human embryonic dental tissues has filled in the blank of cell atlas of early development process of human tooth. Currently available single-cell data on human tooth[12] came from early youth or child volunteers, which inevitably missed progenitor cells at early stage. On the other hand, studies based on mouse embryonic tissues might not be immediately translatable to human research due to profound cross-species difference. Rodent incisor has a relatively short lifespan and routinely lost. Many cell types, including ameloblasts and odontoblasts that secrete mineralized enamel and dentin, respectively, are continuously proliferated from by self-renewal cells located in the proximal part of the tooth[22]. Some of the vital structures of this process, such as the transit-amplifying (T-A) zone, were not found in human dental tissue[23]. Thus, it is desirable to generate the dental cell landscape for both human and mice at the developing stage to achieve intact understanding of human dental generation.

Our result revealed that epithelial cells and mesenchyme cells exhibited profound crosstalk and regulated the developmental process of each other via intercellular signaling networks[24]. In mouse incisor, BMP-SMAD-SHH signals from a transient structure Hartwig's epithelial root sheath[25] interacts with mesenchymal cells and regulated root elongation and epithelium maturation[26]. Epithelium could also regulate the mineralization of nearby dentin though secretion of signal molecules[27], including Wnt and fibroblast growth factors[28]. In line with these observations, previous scRNA intercellular signaling network analysis on postnatal pulp[9] and tooth germ[11] tissues have demonstrated the existence of dense communication. Our work extended these knowledges to prenatal epithelium-mesenchyme developmental process and highlighted more signal pathways like JAG1, SFRP2 and COL4A1.

Our study has limitations. Earlier (<15 gestational weeks) and later (>30 gestational weeks) embryonic healthy tissues are difficult to access due to ethnic concerns, and dental developmental process at these stages could not be covered in the current study. The spatial structures at the third dimension are also not complete in different slicing procedures. Future efforts are desired to fill in these blanks.

In conclusion, we constructed a spatiotemporal cell landscape of human embryonic tooth

germ, depicted developmental trajectory of dental epithelium and mesenchyme, and revealed the biological mechanisms underlying these processes. Our result served as a foundation for future research on dental development and stem cell engineering.

Acknowledgement

This work was supported by Research Fund of Shanghai Tongren Hospital, Shanghai Jiaotong University School of Medicine (No: TRYJ2021JC14) and Tongren Xinxing (TRKYRC-xx202207).

Declaration of Competing Interest

The authors declared no competing interest.

Author contribution

Y.S, W.W and J.Q designed and supervised the study. S.S, J.L, K.G, J.Y, L.H, S.W, Y.Y collected and preprocessed the sample, Y.S curated the data, performed the analysis and drafted the manuscript. All authors read, revised and approved the manuscript.

Reference

1. Yu, T. & Klein, O. D. Molecular and cellular mechanisms of tooth development, homeostasis and repair. *Dev.* **147**, (2020).
2. Balic, A. Concise Review: Cellular and Molecular Mechanisms Regulation of Tooth Initiation. *Stem Cells* vol. 37 26–32 (2019).
3. Thesleff, I. Current understanding of the process of tooth formation: Transfer from the laboratory to the clinic. *Aust. Dent. J.* **59**, 48–54 (2014).
4. Matalová, E., Lungová, V. & Sharpe, P. Development of Tooth and Associated Structures. in *Stem Cell Biology and Tissue Engineering in Dental Sciences* 335–346 (Elsevier Inc., 2015). doi:10.1016/B978-0-12-397157-9.00030-8.
5. Wu, J. *et al.* Single-cell RNA sequencing in oral science: Current awareness and perspectives. *Cell Prolif.* **55**, e13287 (2022).
6. Jing, J. *et al.* Spatiotemporal single-cell regulatory atlas reveals neural crest lineage diversification and cellular function during tooth morphogenesis. *Nat. Commun.* **13**, 1–14 (2022).
7. Krivanek, J. *et al.* Dental cell type atlas reveals stem and differentiated cell types in mouse and human teeth. *Nat. Commun.* **11**, 1–18 (2020).
8. Pagella, P., de Vargas Roditi, L., Stadlinger, B., Moor, A. E. & Mitsiadis, T. A. A single-cell atlas of human teeth. *iScience* **24**, 102405 (2021).
9. Yin, W., Liu, G., Li, J. & Bian, Z. Landscape of Cell Communication in Human Dental Pulp. *Small methods* **5**, (2021).
10. Lee, S. *et al.* Single-Cell RNA Sequencing Analysis of Human Dental Pulp Stem Cell and Human Periodontal Ligament Stem Cell. *J. Endod.* **48**, 240–248 (2022).
11. Shi, Y. *et al.* A single-cell interactome of human tooth germ from growing third molar elucidates signaling networks regulating dental development. *Cell Biosci.* **11**, 1–17 (2021).
12. Hermans, F. *et al.* Establishment of inclusive single-cell transcriptome atlases from mouse and human tooth as powerful resource for dental research. *Front. Cell Dev. Biol.* **10**, 1990 (2022).

13. Wang, B. *et al.* Expression patterns of WNT/β-CATENIN signaling molecules during human tooth development. *J. Mol. Histol.* **45**, 487–496 (2014).
14. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
15. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science (80-.).* **352**, 189–196 (2016).
16. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
17. Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
18. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omi. A J. Integr. Biol.* **16**, 284–287 (2012).
19. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).
20. Tanevski, J., Flores, R. O. R., Gabor, A., Schapiro, D. & Saez-Rodriguez, J. Explainable multiview framework for dissecting spatial relationships from highly multiplexed data. *Genome Biol.* **23**, 1–31 (2022).
21. Liu, F. & Millar, S. E. Wnt/β-catenin signaling in oral tissue development and disease. *J. Dent. Res.* **89**, 318–330 (2010).
22. Seidel, K. *et al.* Resolving stem and progenitor cells in the adult mouse incisor through gene coexpression analysis. *Elife* **6**, (2017).
23. Kuang-Hsien Hu, J., Mushegyan, V. & Klein, O. D. On the cutting edge of organ renewal: Identification, regulation, and evolution of incisor stem cells. *genesis* **52**, 79–92 (2014).
24. Puthiyaveetil, J. S. V., Kota, K., Chakkarayan, R., Chakkarayan, J. & Thodiyil, A. K. P. Epithelial – Mesenchymal Interactions in Tooth Development and the Significant Role of Growth Factors and Genes with Emphasis on Mesenchyme – A Review. *J. Clin. Diagn. Res.* **10**, ZE05 (2016).
25. Huang, X., Xu, X., Bringas, P., Hung, Y. P. & Chai, Y. Smad4-Shh-Nfic signaling cascade-mediated epithelial-mesenchymal interaction is crucial in regulating tooth root development. *J. Bone Miner. Res.* **25**, 1167–1178 (2010).
26. Li, J. *et al.* BMP-SHH Signaling Network Controls Epithelial Stem Cell Fate via Regulation of Its Niche in the Developing Tooth. *Dev. Cell* **33**, 125–135 (2015).
27. Lavicky, J. *et al.* The Development of Dentin Microstructure Is Controlled by the Type of Adjacent Epithelium. *J. Bone Miner. Res.* **37**, 323–339 (2022).
28. Hermans, F., Hemeryck, L., Lambrichts, I., Bronckaers, A. & Vankelecom, H. Intertwined Signaling Pathways Governing Tooth Development: A Give-and-Take Between Canonical Wnt and Shh. *Front. cell Dev. Biol.* **9**, (2021).

Figure Legend

Figure 1 Overview of the human embryonic tooth cell atlas. A: schematic diagram of sample collection. B: UMAP plot of single-cell clustering. SCP: Schwann cell precursor. HSC: hematopoietic stem cells. C: percentage of each cell type and cell cycle phase in samples from

different gestational weeks. Cell cluster color corresponded to B. D: Marker scaled expression of each cluster. E: H-E staining of spatial transcriptome slice. F: Cell type annotation of each spot, defined as the cell type with the highest probability. G: Marker scaled expression of cell markers.

Figure 2 In-depth analysis of dental epithelium. A: UMAP plot and pseudotime analysis of epithelium cell clusters. Arrows marked development direction. Diff amelo: differentiated ameloblast. SC: epithelial stem cell. SI: striatum intermedium. B-C: similar to Figure 1 C-D, but for epithelium clusters. D: Module expression along pseudotime. For each branch, we identified an up-regulation module and a down-regulation module. E: GO term enriched for each module. ECM：extracellular matrix. F: Transcription factors whose targets were enriched for module genes. Color denoted the module that were enriched, transparency showed enrichment p-value, font size showed odds ratio. G: Regulon expression of selected TF.

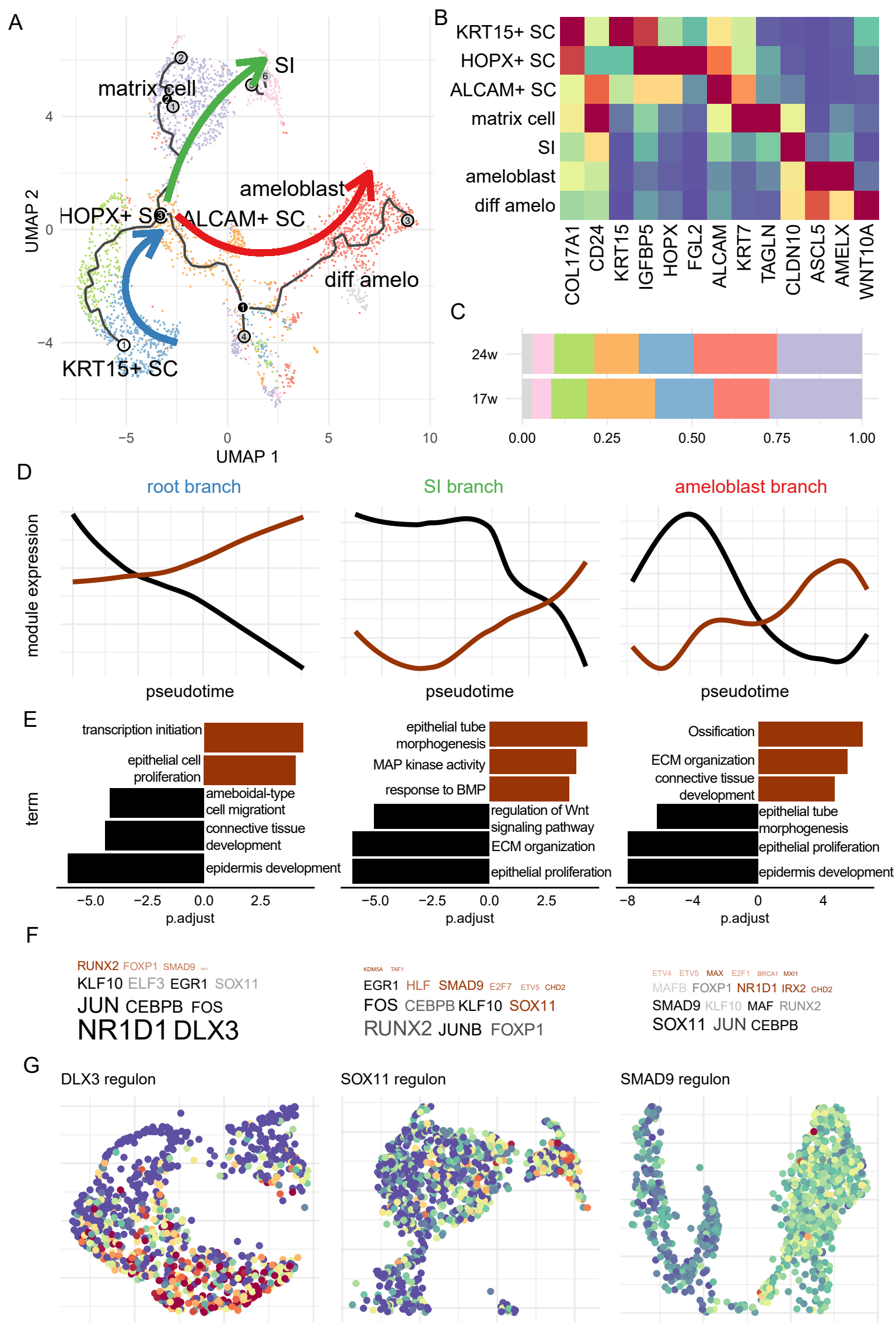
Figure 3 Spatial transcriptome of dental epithelium. A: Cell type annotation of each spot. B: Cell type probability spatial distribution. C: Cell type probability distribution per developmental stage. D: misty result of spatial adjacency, quantified by model importance score. RMSE: Root mean squared error of each model.

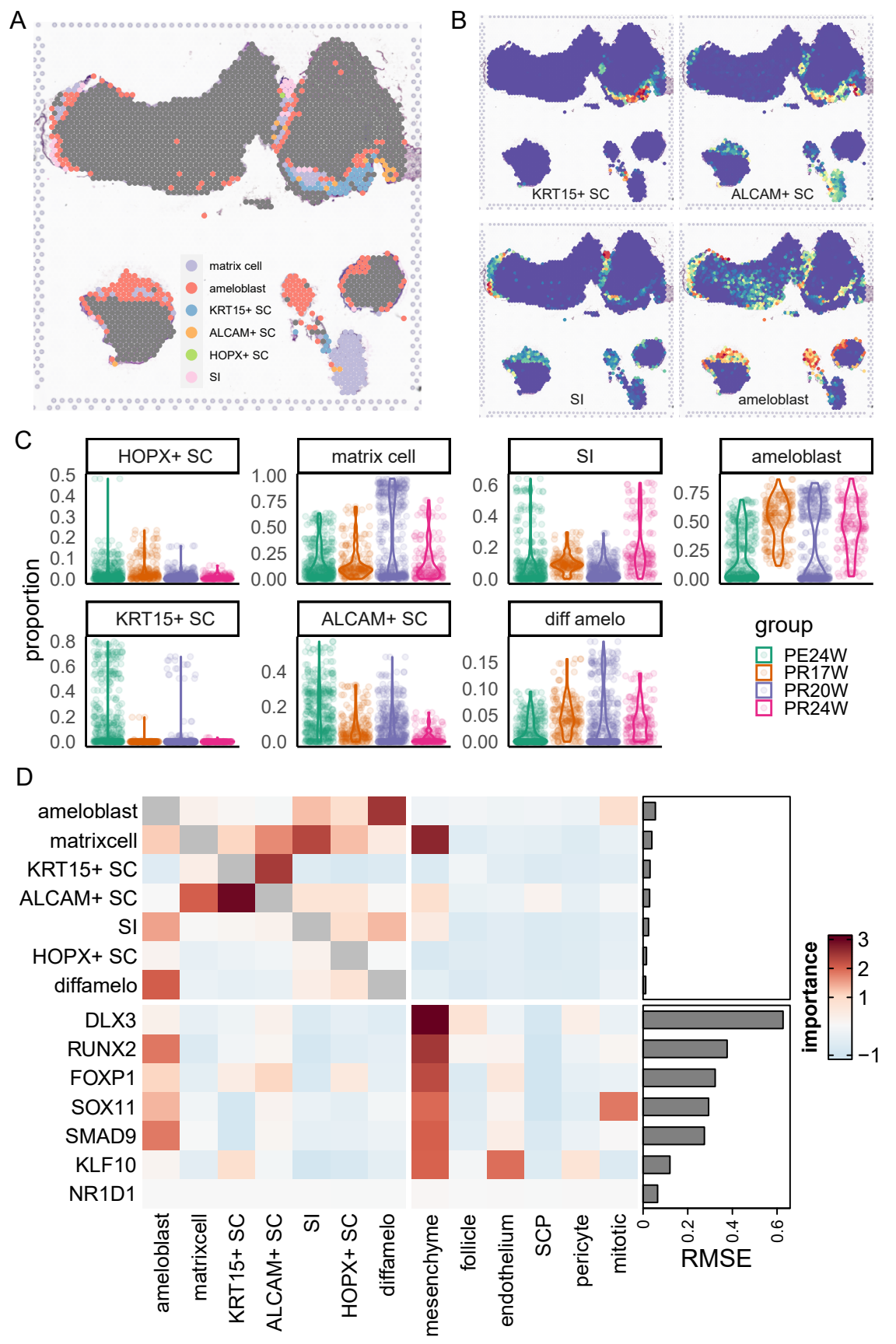
Figure 4 In-depth analysis of dental mesenchyme. A -C: similar to Figure 2 A-C, but for epithelium clusters. Pseudotime trajectory was shown in Figure S7. OD: pre-odontoblast. AP: apical progenitor. MSC: mesenchymal stem cell. D: Expression of four waves of gene along mesenchyme development. E: Heatmap of module gene expression along pseudotime. Yellow bar denoted that the gene was predicted to be target of the corresponding transcription factor. Word cloud showed GO term enriched by each wave. Font size denoted enrichment odds ratio, color denoted enrichment p value. F: Expression of selected regulons.

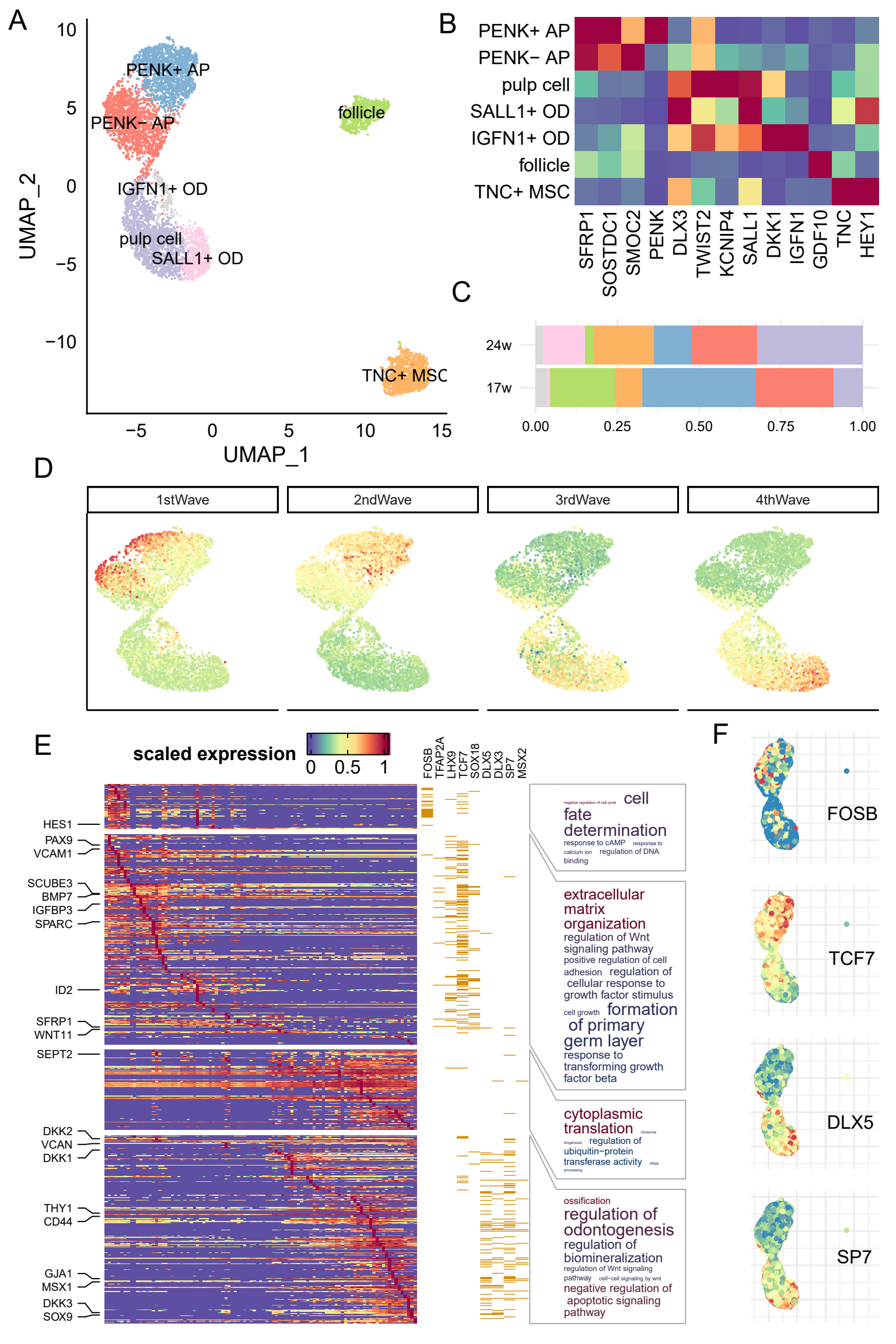
Figure 5 Spatial transcriptome of dental mesenchyme. A: Cell type annotation of each spot. B: Cell type probability spatial distribution and marker gene spatial expression. C: Cell type probability distribution per developmental stage. D: misty result of spatial adjacency, quantified by model importance score. RMSE: Root mean squared error of each model.

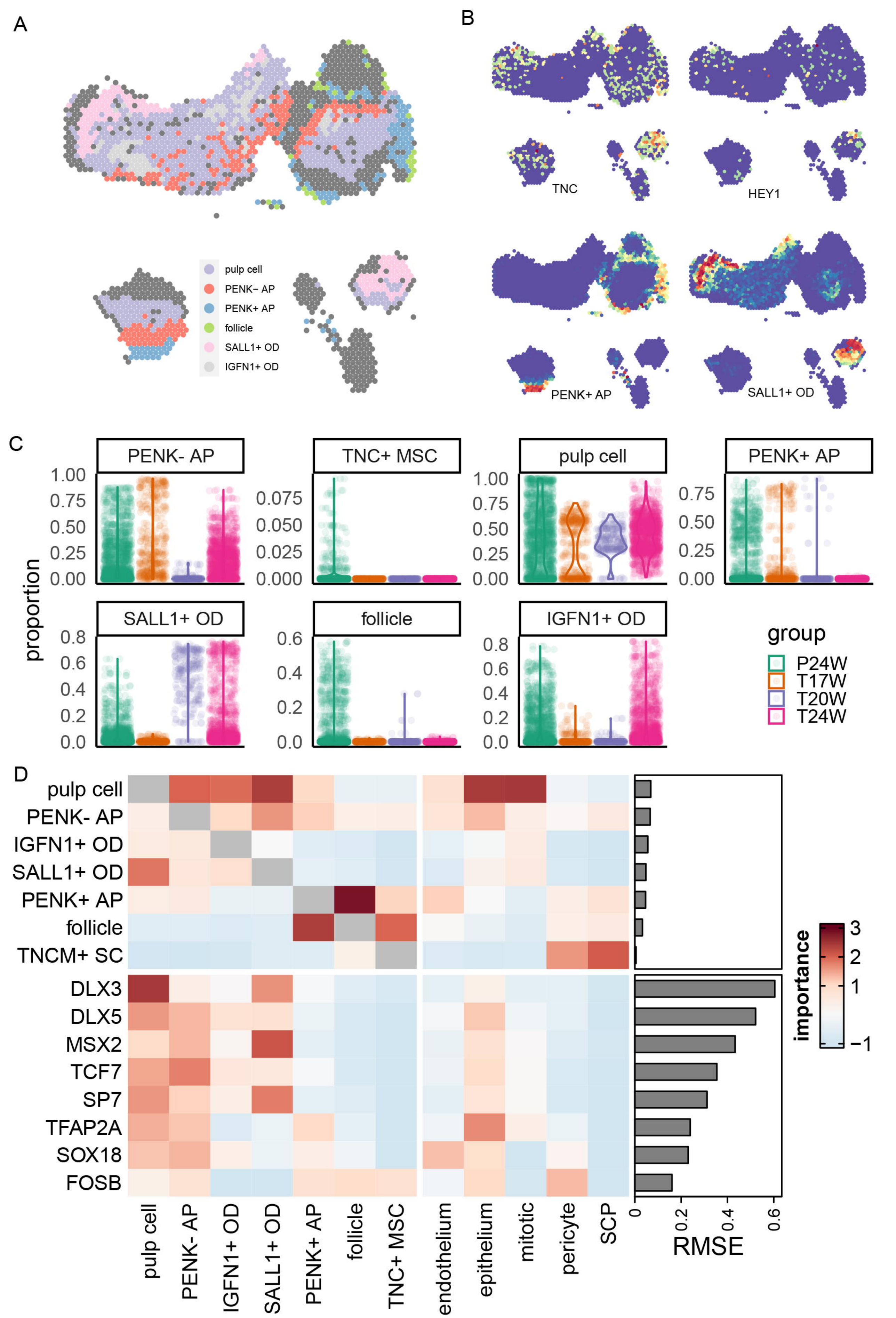
Figure 6 Signaling network regulating epithelium development. A: Each line linked a ligand and a predicted target; line color showed the cell type in which the ligand had the highest expression (corresponded to color bar of Figure 1B). Heatmap showed cell type average expression of ligands. black bar denoted the transcription factor that was predicted to regulate the corresponding target gene. Link weight was calculated by nichenetr. B: Spatial adjacency between ligand and target, quantified by misty model importance score. C: spatial expression of selected ligand and target.

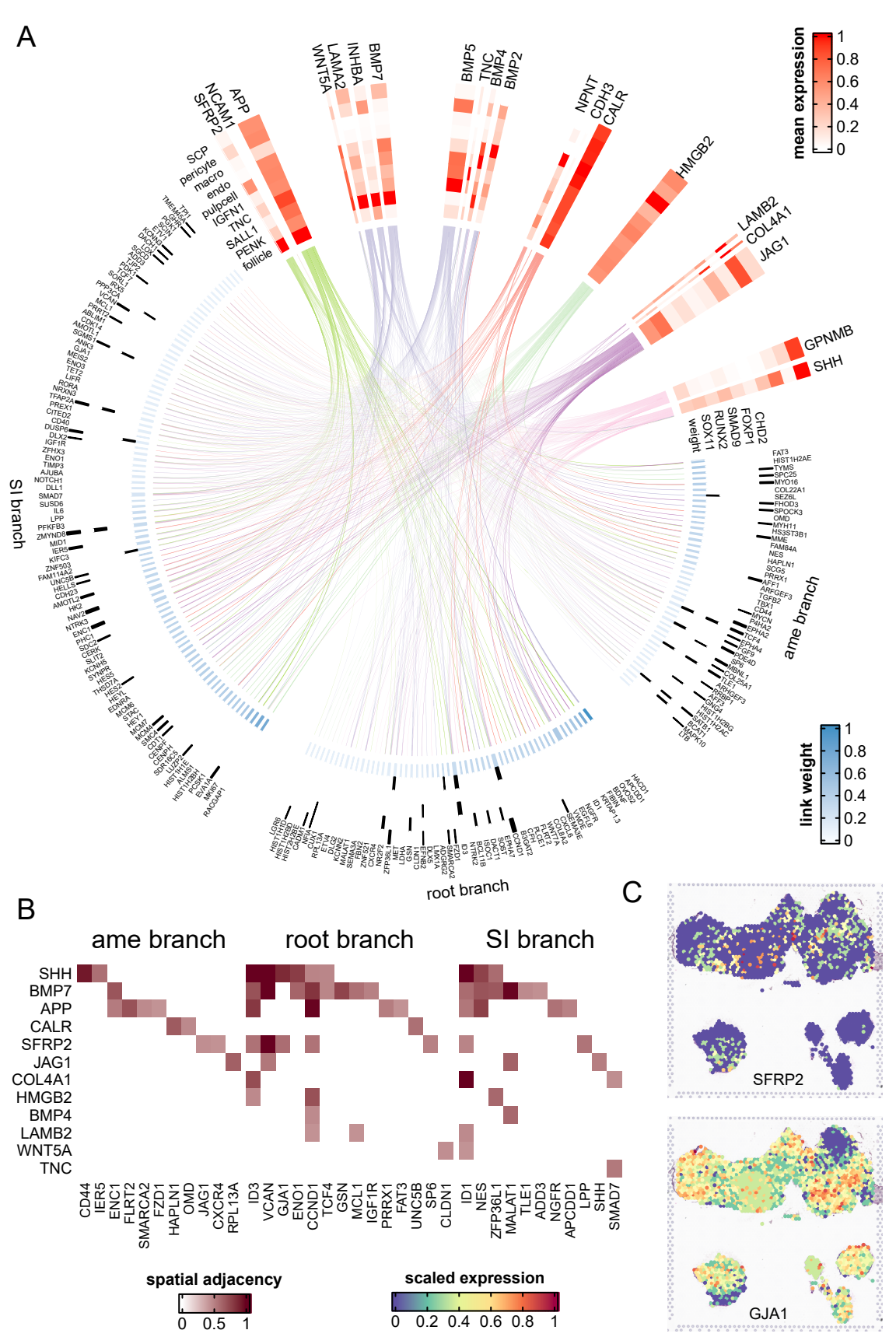
Figure 7 Signaling network regulating mesenchyme development. Similar to Figure 6, but for mesenchyme.

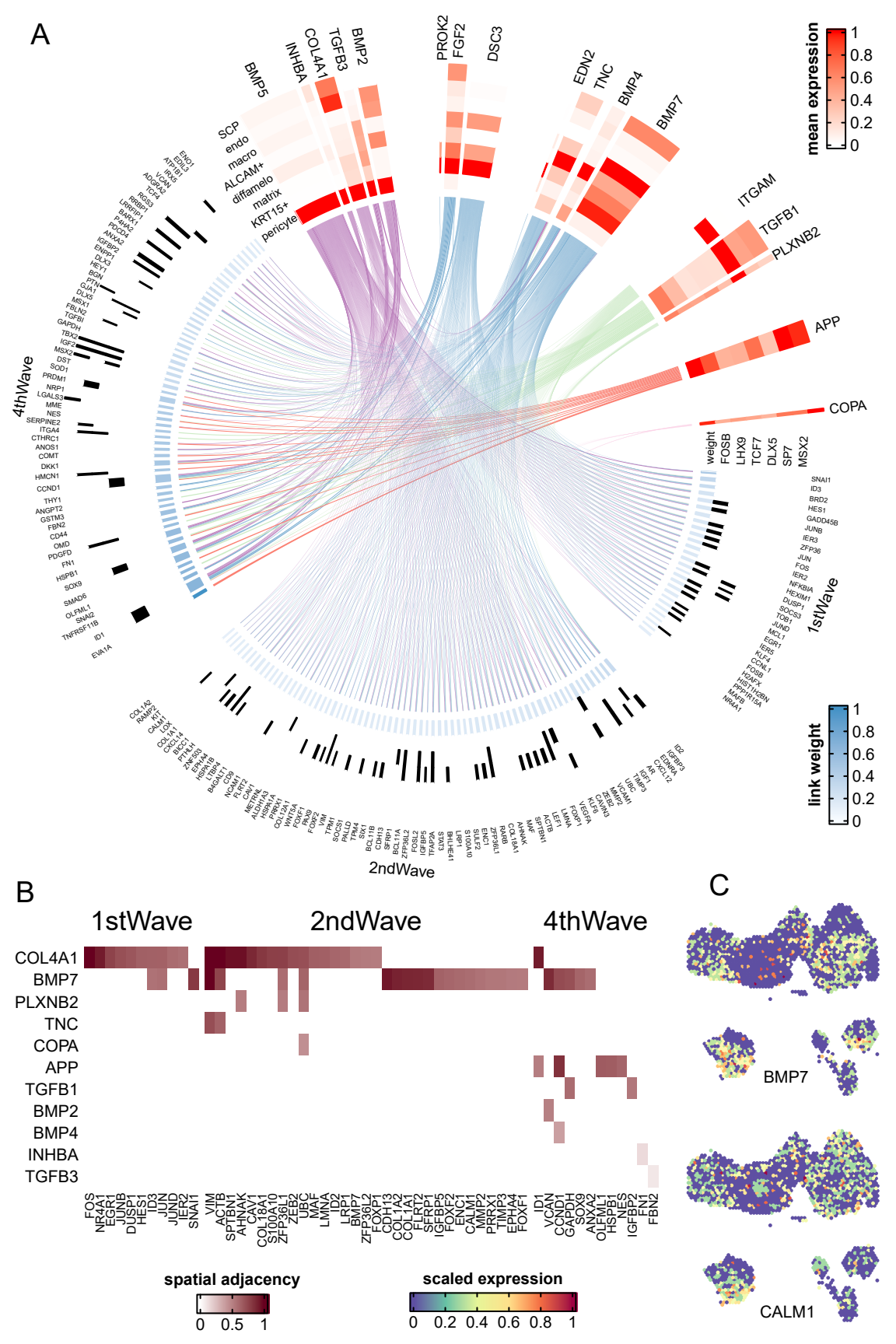

A
matrix cell
SI
ameloblast
HOPX+ SC
ALCAM+ SC
diff amelo
KRT15+ SC
UMAP 2
UMAP 1
B
KRT15+ SC
HOPX+ SC
ALCAM+ SC
matrix cell
SI
ameloblast
diff amelo
COL17A1
CD24
KRT15
IGFBP5
HOPX
FGL2
ALCAM
KRT7
TAGLN
CLDN10
ASCL5
AMELX
WNT10A
C
24w
17w
D
root branch
SI branch
ameloblast branch
module expression
pseudotime
E
term
transcription initiation
epithelial cell proliferation
ameboidal-type cell migrationt
connective tissue development
epidermis development
epithelial tube morphogenesis
MAP kinase activity
response to BMP
regulation of Wnt signaling pathway
ECM organization
epithelial proliferation
Ossification
ECM organization
connective tissue development
epithelial tube morphogenesis
epithelial proliferation
epidermis development
p.adjust
F
RUNX2 FOXP1 SMAD9
KLF10 ELF3 EGR1 SOX11
JUN CEBPB FOS
NR1D1 DLX3
KDM5A TAF1
EGR1 HLF SMAD9 E2F7 ETV5 CHD2
FOS CEBPB KLF10 SOX11
RUNX2 JUNB FOXP1
ETV4 ETV5 MAX E2F1 BRCA1 MXI1
MAFB FOXP1 NR1D1 IRX2 CHD2
SMAD9 KLF10 MAF RUNX2
SOX11 JUN CEBPB
G
DLX3 regulon
SOX11 regulon
SMAD9 regulon


A
matrix cell
ameloblast
KRT15+ SC
ALCAM+ SC
HOPX+ SC
SI
B
KRT15+ SC
ALCAM+ SC
SI
ameloblast
C
HOPX+ SC
matrix cell
SI
ameloblast
KRT15+ SC
ALCAM+ SC
diff amelo
proportion
group
PE24W
PR17W
PR20W
PR24W
D
ameloblast
matrixcell
KRT15+ SC
ALCAM+ SC
SI
HOPX+ SC
diffamelo
DLX3
RUNX2
FOXP1
SOX11
SMAD9
KLF10
NR1D1
ameloblast
matrixcell
KRT15+ SC
ALCAM+ SC
SI
HOPX+ SC
diffamelo
mesenchyme
follicle
endothelium
SCP
pericyte
mitotic
RMSE
importance

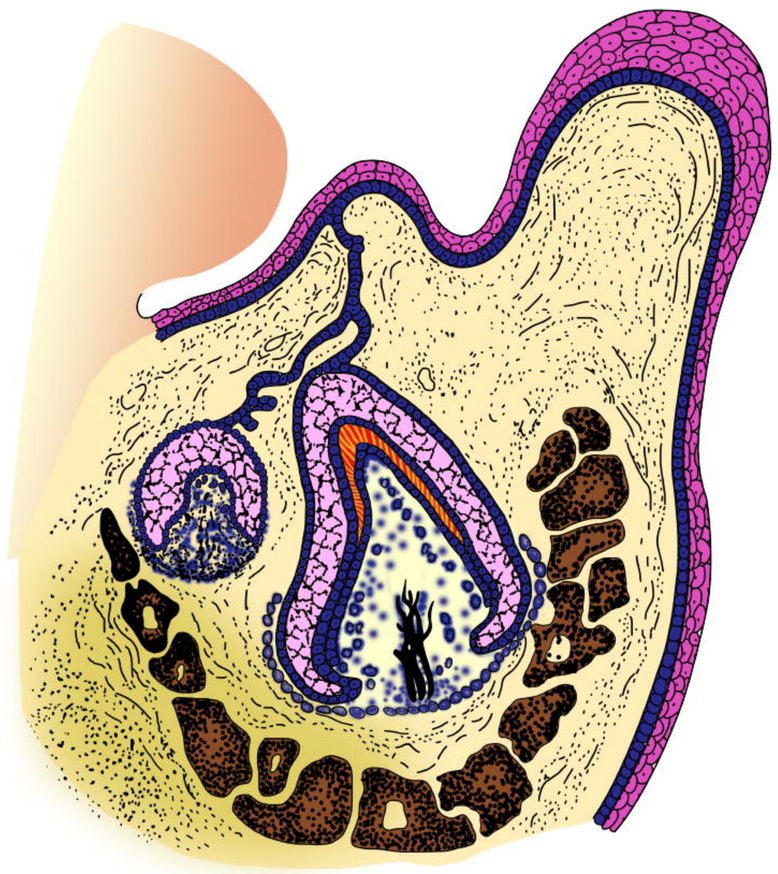

A
PENK+ AP
PENK− AP
follicle
IGFN1+ OD
pulp cell
SALL1+ OD
TNC+ MSC
UMAP_2
UMAP_1
B
PENK+ AP
PENK− AP
pulp cell
SALL1+ OD
IGFN1+ OD
follicle
TNC+ MSC
SFRP1
SOSTDC1
SMOC2
PENK
DLX3
TWIST2
KCNIP4
SALL1
DKK1
IGFN1
GDF10
TNC
HEY1
C
24w
17w
0.00
0.25
0.50
0.75
1.00
D
1stWave
2ndWave
3rdWave
4thWave
E
scaled expression
0
0.5
1
FOSB
TFAP2A
LHX9
TCF7
SOX18
DLX5
DLX3
SP7
MSX2
HES1
PAX9
VCAM1
SCUBE3
BMP7
IGFBP3
SPARC
ID2
SFRP1
WNT11
SEPT2
DKK2
VCAN
DKK1
THY1
CD44
GJA1
MSX1
DKK3
SOX9
cell fate determination
response to cAMP
response to calcium ion
regulation of DNA binding
extracellular matrix organization
regulation of Wnt signaling pathway
positive regulation of cell adhesion
regulation of cellular response to growth factor stimulus
formation of primary germ layer
response to transforming growth factor beta
cytoplasmic translation
regulation of ubiquitin−protein transferase activity
ossification
regulation of odontogenesis
regulation of biomineralization
regulation of Wnt signaling pathway
negative regulation of apoptotic signaling pathway
F
FOSB
TCF7
DLX5
SP7


A
pulp cell
PENK− AP
PENK+ AP
follicle
SALL1+ OD
IGFN1+ OD
B
TNC
HEY1
PENK+ AP
SALL1+ OD
C
PENK- AP
TNC+ MSC
pulp cell
PENK+ AP
SALL1+ OD
follicle
IGFN1+ OD
proportion
group
P24W
T17W
T20W
T24W
D
pulp cell
PENK- AP
IGFN1+ OD
SALL1+ OD
PENK+ AP
follicle
TNCM+ SC
DLX3
DLX5
MSX2
TCF7
SP7
TFAP2A
SOX18
FOSB
pulp cell
PENK- AP
IGFN1+ OD
SALL1+ OD
PENK+ AP
follicle
TNC+ MSC
endothelium
epithelium
mitotic
pericyte
SCP
RMSE
importance

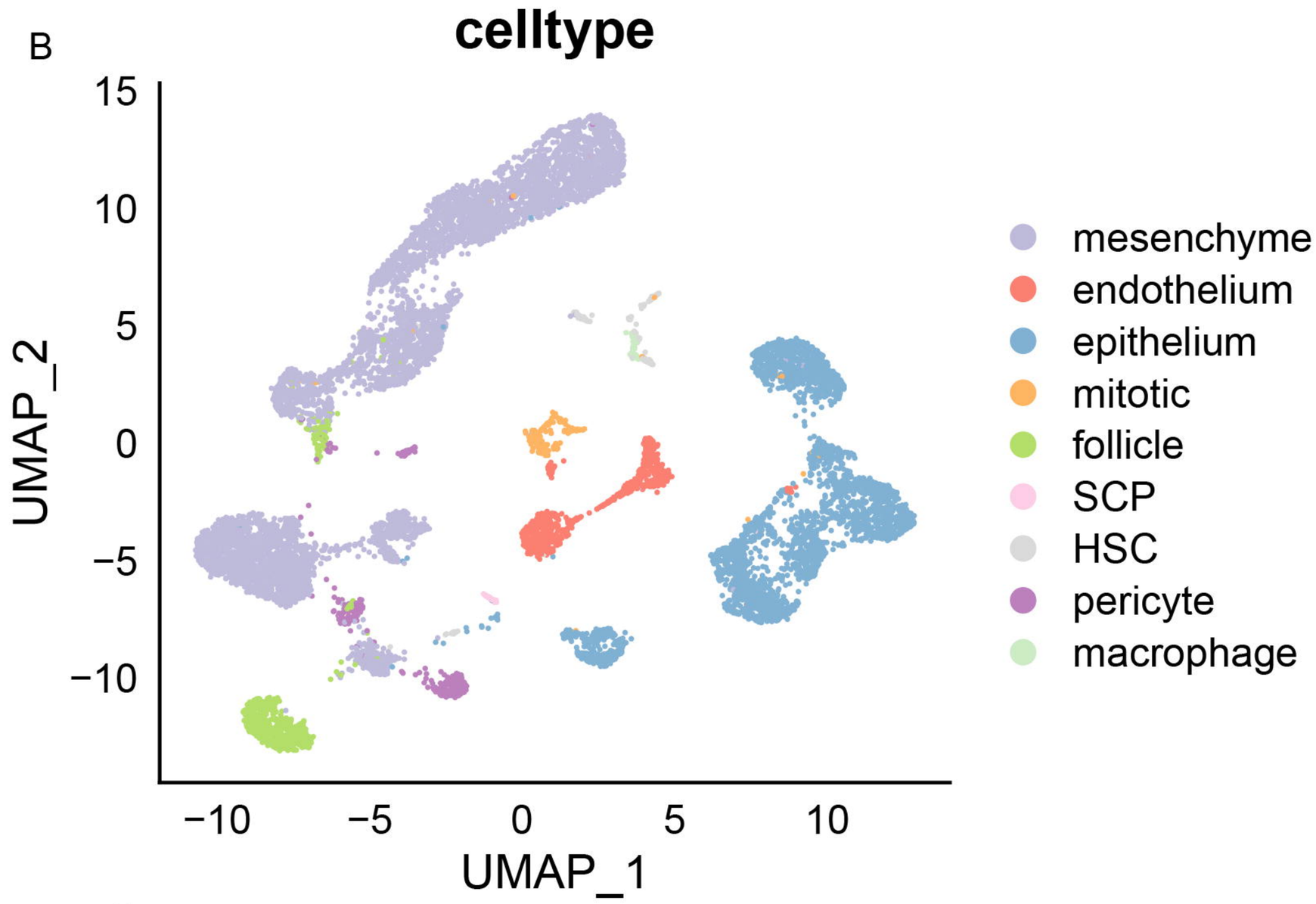

A
mean expression
1
0.8
0.6
0.4
0.2
0
APP
NCAM1
SFRP2
WNT5A
LAMA2
INHBA
BMP7
BMP5
TNC
BMP4
BMP2
NPNT
CDH3
CALR
HMGB2
LAMB2
COL4A1
JAG1
GPNMB
SHH
CHD2
FOXP1
SMAD9
RUNX2
SOX11
weight
SCP
pericyte
macro
endo
pulpcell
IGFN1
TNC
SALL1
PENK
follicle
SI branch
ame branch
root branch
link weight
B
ame branch
root branch
SI branch
SHH
BMP7
APP
CALR
SFRP2
JAG1
COL4A1
HMGB2
BMP4
LAMB2
WNT5A
TNC
CD44
IER5
ENC1
FLRT2
SMARCA2
FZD1
HAPLN1
OMD
JAG1
CXCR4
RPL13A
ID3
VCAN
GJA1
ENO1
CCND1
TCF4
GSN
MCL1
IGF1R
PRRX1
FAT3
UNC5B
SP6
CLDN1
ID1
NES
ZFP36L1
MALAT1
TLE1
ADD3
NGFR
APCDD1
LPP
SHH
SMAD7
spatial adjacency
0
0.5
1
scaled expression
0
0.2
0.4
0.6
0.8
1
C
SFRP2
GJA1


A
mean expression
1
0.8
0.6
0.4
0.2
0
BMP5
INHBA
COL4A1
TGFB3
BMP2
PROK2
FGF2
DSC3
EDN2
TNC
BMP4
BMP7
ITGAM
TGFB1
PLXNB2
APP
COPA
SCP
endo
macro
ALCAM+
diffamelo
matrix
KRT15+
pericyte
4thWave
1stWave
2ndWave
weight
FOSB
LHX9
TCF7
DLX5
SP7
MSX2
link weight
1
0.8
0.6
0.4
0.2
0
B
1stWave
2ndWave
4thWave
COL4A1
BMP7
PLXNB2
TNC
COPA
APP
TGFB1
BMP2
BMP4
INHBA
TGFB3
spatial adjacency
0
0.5
1
scaled expression
0
0.2
0.4
0.6
0.8
1
C
BMP7
CALM1

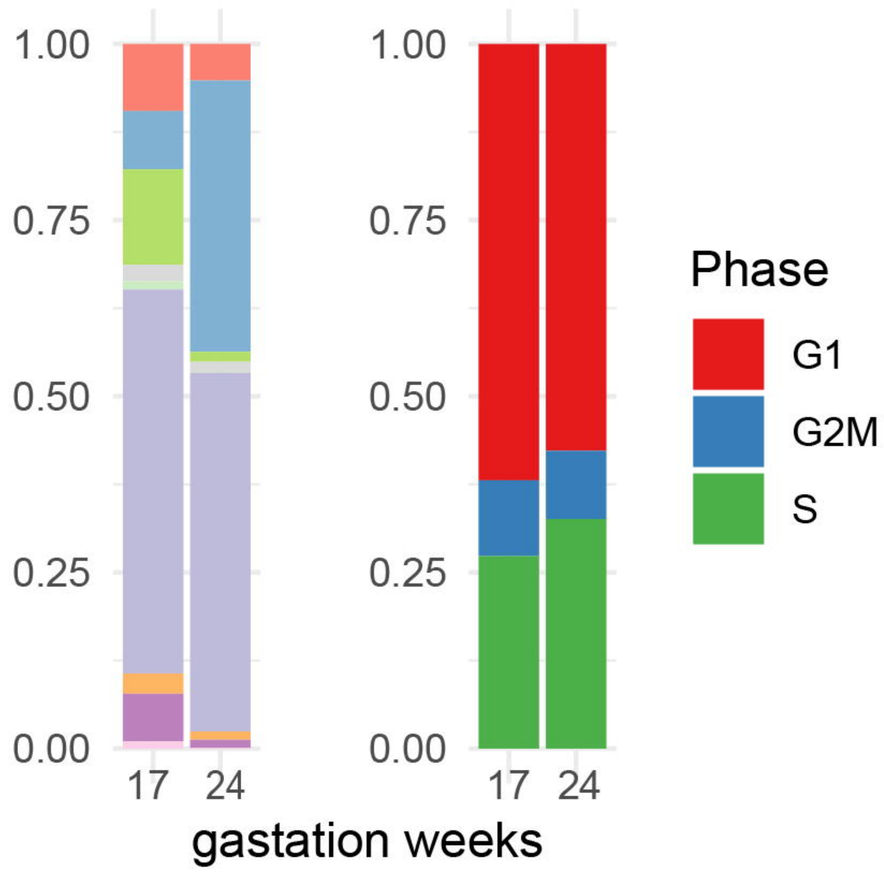
A



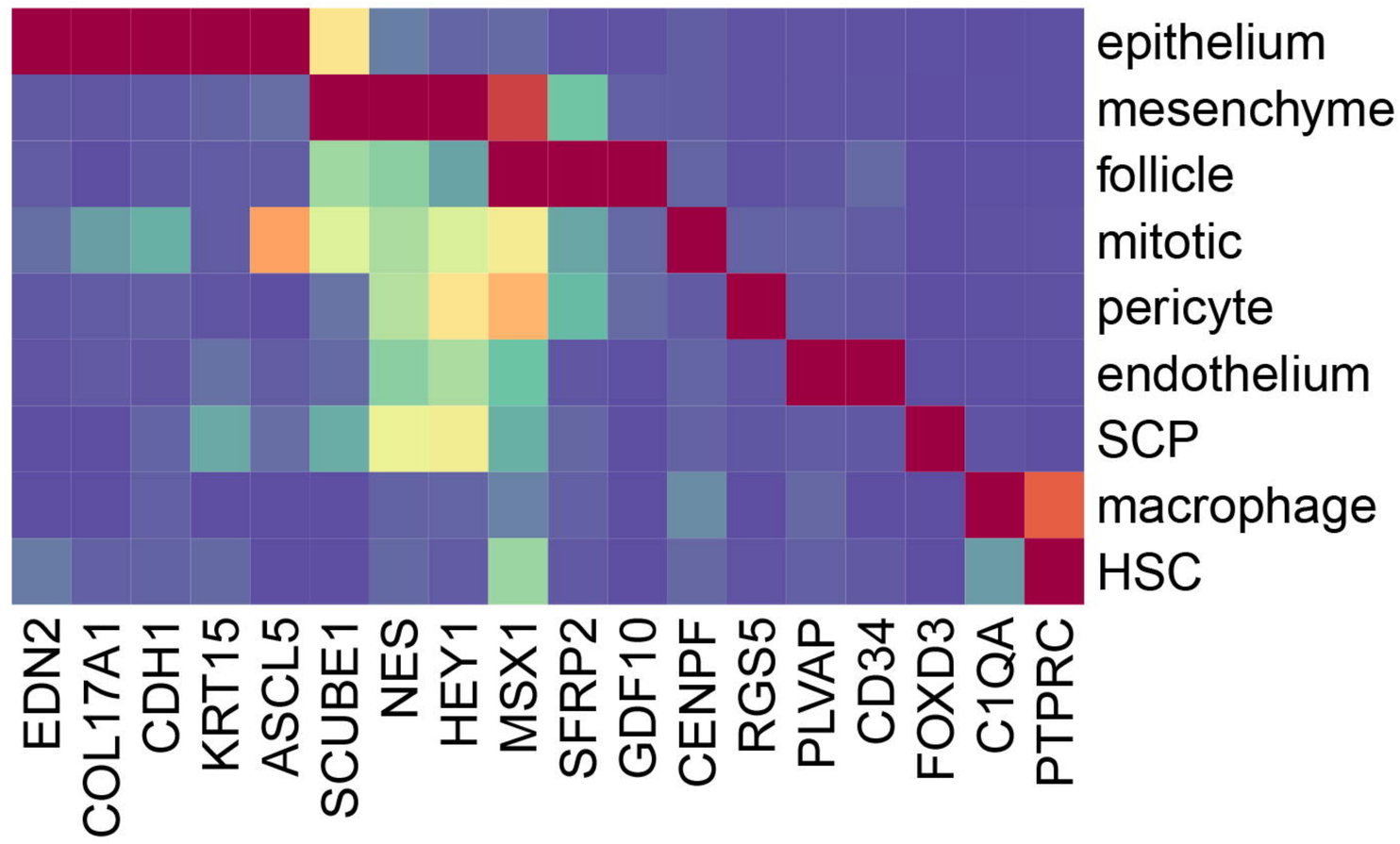
B


celltype
UMAP_2
UMAP_1
15
10
5
0
−5
−10
−10
−5
0
5
10
mesenchyme
endothelium
epithelium
mitotic
follicle
SCP
HSC
pericyte
macrophage

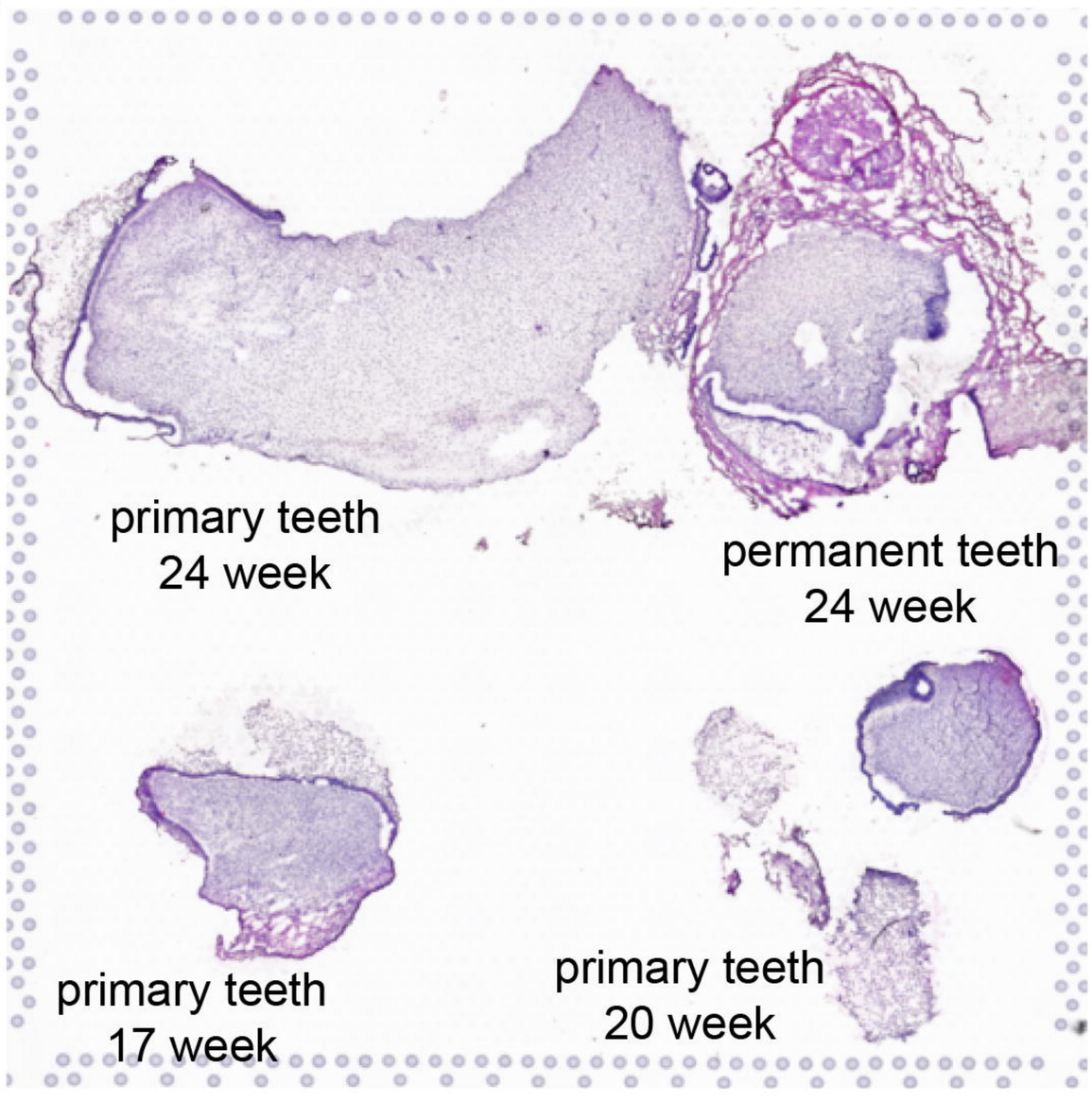
C


1.00
0.75
0.50
0.25
0.00
17
24
gastation weeks
Phase
G1
G2M
S

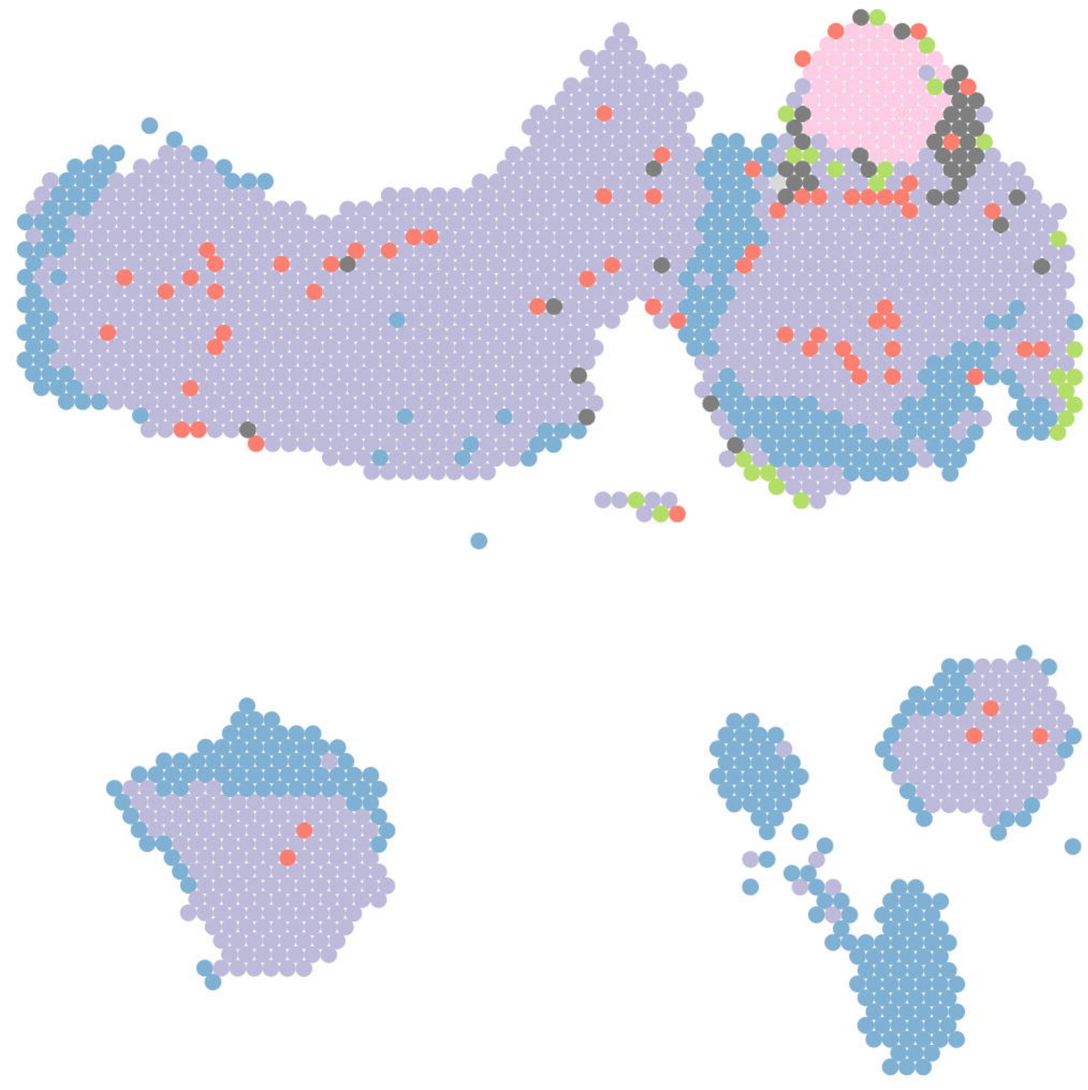
D


epithelium
mesenchyme
follicle
mitotic
pericyte
endothelium
SCP
macrophage
HSC
EDN2
COL17A1
CDH1
KRT15
ASCL5
SCUBE1
NES
HEY1
MSX1
SFRP2
GDF10
CENPF
RGS5
PLVAP
CD34
FOXD3
C1QA
PTPRC

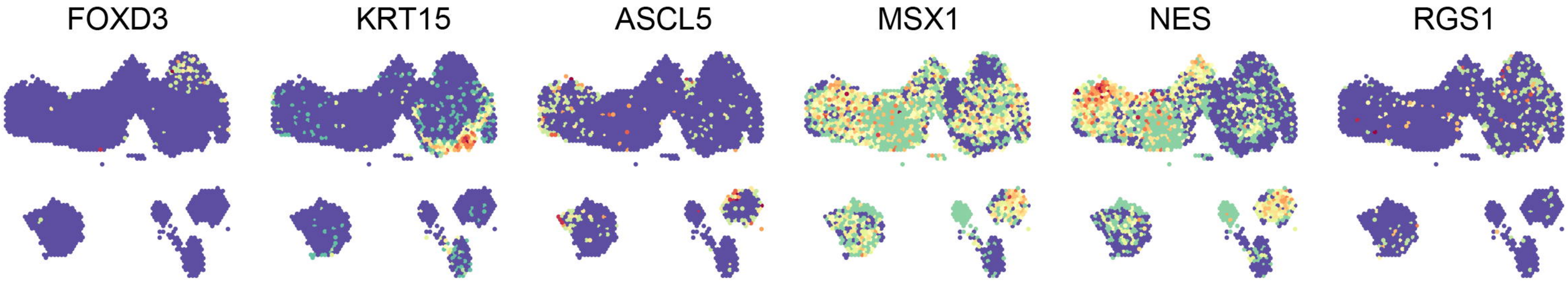
E


primary teeth
24 week
permanent teeth
24 week
primary teeth
17 week
primary teeth
20 week

F



G


FOXD3
KRT15
ASCL5
MSX1
NES
RGS1