ORIGINAL RESEARCH

Single-cell Transcriptomics Reveals Early Molecular and Immune Alterations Underlying the Serrated Neoplasia Pathway Toward Colorectal Cancer

Yu-Jie Zhou,[1,*] Xiao-Fan Lu,[2,*] Huimin Chen,[1,*] Xin-Yuan Wang,[1] Wenxuan Cheng,[2] Qing-Wei Zhang,[1] Jin-Nan Chen,[1] Xiao-Yi Wang,[1] Jing-Zheng Jin,[1] Fang-Rong Yan,[2] Haoyan Chen,[3] and Xiao-Bo Li[1]

[1]Division of Gastroenterology and Hepatology, Key Laboratory of Gastroenterology and Hepatology, Ministry of Health, Shanghai Institute of Digestive Disease, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China; [2]State Key Laboratory of Natural Medicines, Research Center of Biostatistics and Computational Pharmacy, China Pharmaceutical University, Nanjing, China; and [3]State Key Laboratory for Oncogenes and Related Genes, Shanghai Cancer Institute, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

SUMMARY

The study identified epithelial cells specific to serrated lesions (SLs) with dysregulated expression of genes related to redox balance, and changes specific to SLs in non-parenchymal compartment (eg, accumulation of tissue-resident memory T cells and PDGFRA[+] fibroblasts) by depicting the single-cell landscape of 16 SLs.

BACKGROUND & AIMS: Approximately one-third of colorectal cancers develop from serrated lesions (SLs), including hyperplastic polyp (HP), sessile serrated lesion (SSL), traditional serrated adenoma (TSA), and SSL with dysplasia (SSLD) through the serrated neoplasia pathway, which progresses faster than the conventional adenoma-carcinoma pathway. We sought to depict the currently unclarified molecular and immune alterations by the single-cell landscape in SLs.

METHODS: We performed single-cell RNA sequencing of 16 SLs (including 4 proximal HPs, 5 SSLs, 2 SSLDs, and 5 TSAs) vs 3 normal colonic tissues.

RESULTS: A total of 60,568 high-quality cells were obtained. Two distinct epithelial clusters with redox imbalance in SLs were observed, along with upregulation of tumor-promoting SerpinB6 that regulated ROS level. Epithelial clusters of SSL and TSA showed distinct molecular features: SSL-specific epithelium manifested overexpressed proliferative markers with Notch pathway activation, whereas TSA-specific epithelium showed Paneth cell metaplasia with aberrant lysozyme expression. As for immune contexture, enhanced cytotoxic activity of CD8[+] T cells was observed in SLs; it was mainly attributable to increased proportion of CD103[+]CD8[+] tissue-resident memory T cells, which might be regulated by retinoic acid metabolism. Microenvironment of SLs was generally immune-activated, whereas some immunosuppressive cells (regulatory T cells, anti-inflammatory macrophages, MDK[+]IgA[+] plasma cells, MMP11-secreting PDGFRA[+] fibroblasts) also emerged at early stage and further accumulated in SSLD.

CONCLUSION: Epithelial, immune, and stromal components in the serrated pathway undergo fundamental alterations. Future

molecular subtypes of SLs and potential immune therapy might be developed.

Colorectal cancer (CRC) ranks as the third most common malignant tumor and the third leading cause of cancer deaths around the globe.[1] CRC originates from premalignant lesions through 2 major pathways: conventional adenoma-carcinoma sequence and alternative serrated tumorigenesis pathway. Characterized by saw-toothed appearance of crypts, serrated lesions (SLs) account for precancerous lesions of 15% to 35% sporadic CRC.[2] According to the latest World Health Organziation classification, SLs consist of the following categories based

 Most current article

on pathological features: hyperplastic polyp (HP), sessile serrated lesion (SSL), SSL with dysplasia (SSLD), traditional serrated adenoma (TSA), and unclassified.[3] Proximal (right-sided) microvesicular HPs are precursors of SSLs. It is now recognized that SSLs and TSAs have the malignant potential to develop into serrated CRC.[4]

SSL locates in proximal colon with predominant BRAF[V600E] mutation, which can progress into BRAF-mutant right-sided CRC with either microsatellite instability-high (MSI-H) or microsatellite stable (MSS).[5] TSA locates in distal colon with BRAF or KRAS mutation, which is thought to develop into BRAF/KRAS-mutant serrated CRC with MSS.[6] Of note, some case studies revealed that lesions arising from serrated pathway showed accelerated carcinogenesis, with reports of rapid transformation from SSL to invasive CRC within several months,[7,8] although there was no definite evidence,[9] and contradictory opinions existed.[10] Furthermore, SSL is easily missed during colonoscopy examination due to its flat appearance. As such, SSLs have become a major cause of interval CRCs after a negative colonoscopy.[11]

Since the concept of the serrated pathway was first proposed just 3 decades ago,[12] its major molecular alterations, and its interaction with the immunological environment are not well-characterized. Understanding the pathogenesis of serrated neoplasia is vital for precision treatment and early intervention for CRC. However, a key challenge in further elucidating the serrated pathway is that it is difficult to clearly distinguish serrated pathway-derived invasive CRC from those originated from conventional adenomas, owing to the loss of serrated morphology at the late stage.[13] To fulfill this gap, in this study, we focused on well-characterized early lesions of the serrated subtypes and performed single-cell RNA sequencing (scRNA-seq) of 16 SL samples, for more in-depth understanding of pathogenesis of serrated colorectal neoplasia and the serrated pathway to CRC.

Results

Single-cell Transcriptomic Atlas of Human SLs

To explore transcriptional alterations in serrated tumorigenesis, we generated an scRNA-seq profile from 3 normal contrasts (NCs), 4 right-sided microvesicular HPs, 5 SSLs, 2 SSLDs, and 5 TSAs samples using $10\times$ Genomics sequencing (Figure 1, A). The clinical characteristics and endoscopic imaging of these patients are shown in Supplementary Table 1 and Figure 1, B–E. All serrated lesions used for analysis used in our study were sporadic.

After human genome mapping and quality filtering, we obtained a total of 60,568 high-quality cells, of which 11,661 cells were from NC, 12,465 cells were from HP, 16,713 cells were from SSL, 3003 cells were from SSLD, and 16,726 cells were from TSA (Figure 1, F–H; Figure 2, A).

We identified 16 clusters in total through Louvain clustering and uniform manifold approximation and projection (UMAP) for dimensionality reduction and visualization (Figure 2, B). We then reclassified these Louvain clusters into global cellular compartments using graph-based clustering and known classic markers (epithelial cells: EPCAM; T cells: CD3D; B cells: MS4A1, CD79A; myeloid cells: CD14, CD163; stromal cells: COL1A1, PECAM1). Mast cells were removed from analysis as fewer than 1000 cells were detected. Each cell type was from multiple patients, which indicated correction of the batch effect (Figure 2, C–D). The top 10 differentially expressed genes (DEGs; adjusted $P < .05$; $|\log_2FC| > 0.25$) between major cell types further confirmed the clustering results (Supplementary Table 2 and Figure 2, E).

Identification of an Epithelial Cell Population Featuring Redox Imbalance Specifically Existing in SLs

In this study, we identified 20 main epithelial cell types in serrated polyps, after the reclustering of 7294 epithelial cells (Figure 3, A–C and Supplementary Table 3). Seven types of normal cells in large intestine were designated: BEST4[+] colonocytes (high expression of BEST4, OPOT2, LYPD8, CA4, CA7), mature colonocytes (BEST4[lo] GUCA2B[hi], and SLC26A3[hi]), goblet cells (high expression of TFF3, SPINK1, REG4, AGR2, MUC2, and WFDC2), transit amplifying cells (high expression of mitochondrial ATP synthase component ATP5MC1, OLFM4[+], and PCNA[+]), intermediate cells (expressed both absorptive and secretory markers and MKI67), tuff cells (HPGDS[+] and PTGS1[+]), and enteroendocrine cells (SCGN[+] and PCSK1N[+]).

In SLs, we observed a decreased proportion of these normal cell components, with lower expression of endogenous hormones guanylin (GUCA2A) and uroguanylin (GUCA2B), which were involved in intestinal fluid homeostasis and normal secreting function (Figure 3, D). BEST4/OTOP2 cells were a cluster of newly identified pH-sensing absorptive colonocyte expressing calcium-sensitive chloride channel BEST4,[14] which was almost absent in SLs, suggesting impaired ion secretion and intestinal barrier function.

Interestingly, a distinct epithelial subpopulation was found predominantly in SLs but was rarely present in the normal tissue (Figure 3, E). Thus, we named these SL-specific epithelia as Epi-SL (cluster 1, 2, 5, 6, 9, 12, 15, 16, 17). We compared DEGs between Epi-SL and normal colonocytes (BEST4[+] colonocytes, mature colonocytes, and goblet cells). As shown in Figure 3, F–G, top overexpressed genes of Epi-SL included OLFM4 (intestinal stem cell marker), GSTP1 (an antioxidant enzyme), PKM (a key rate-limiting enzyme in glycolysis), and lysozyme (LYZ). Interestingly, LYZ was a known marker of Paneth cell, which predominantly existed in the small intestine and was nearly absent in normal large intestine epithelium.[15] Of note, several significantly upregulated genes in Epi-SL, such as SOX9, MYC, GOLM1, LCN2, S100A11, FOXQ1, ANXA2, ENO1, and PKM (Supplementary Table 4), have been established to play oncogenic roles in colon tumorigenesis,[16-18] indicating the carcinogenic nature of Epi-SL cluster.

To further characterize Epi-SL cells, which represent altered epithelium in premalignant colon SLs, we focused on genes significantly upregulated in this subpopulation

compared with normal absorptive and secretory colonocytes (adjusted $P < .05$; $\log_2 FC > 0.25$). Gene Ontology (GO) enrichment analysis showed that pathways related to response to oxidative stress and oxidoreductase activity were significantly enriched in Epi-SL cluster (all adjusted $P < .05$) (Figure 4, A). These findings suggested that redox
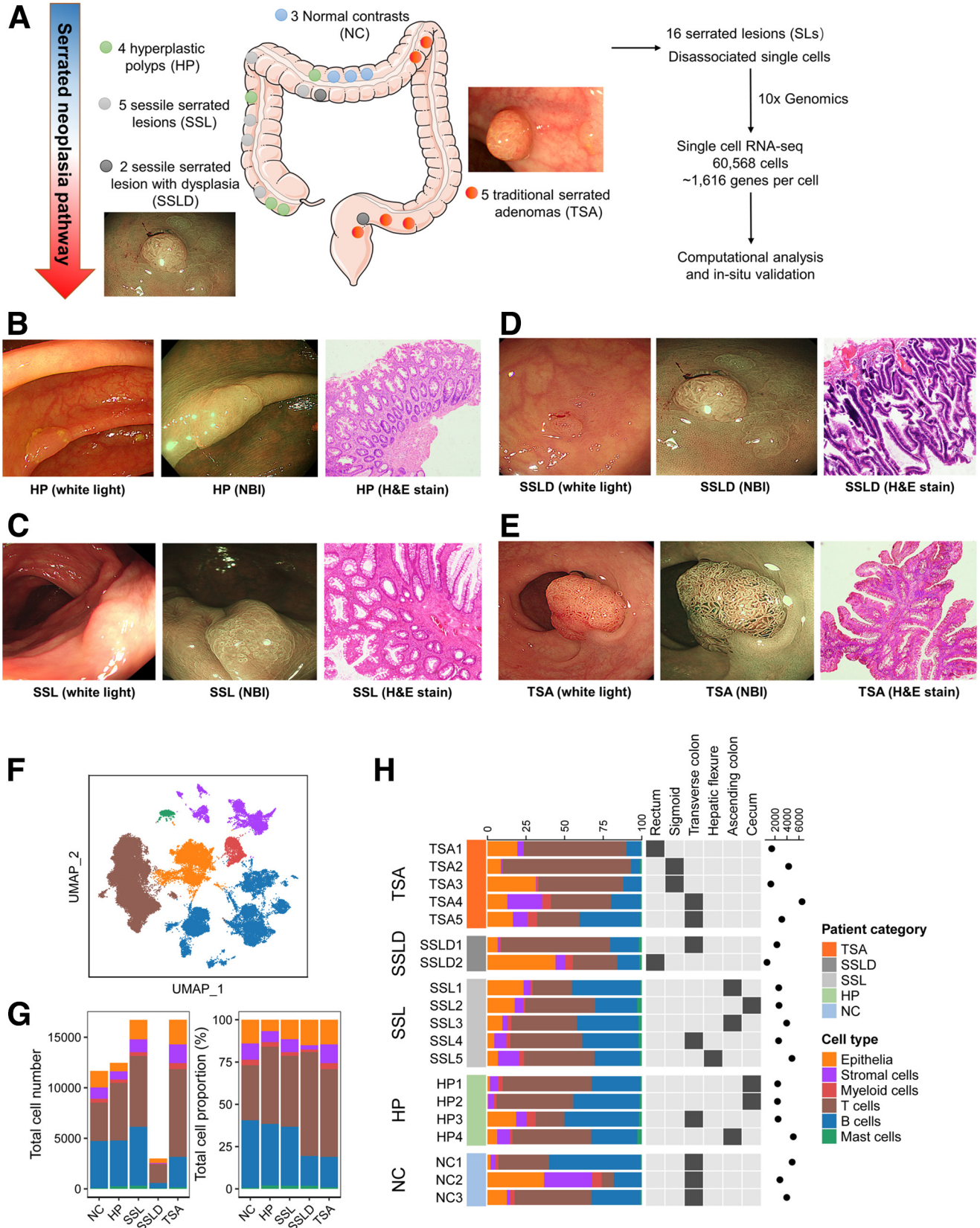
imbalance might play a role in serrated tumorigenesis. In Epi-SL, upregulation of both oxidant (DUOX1, DUOX2, DUOXA1, DUOXA2, GPX2, AQP5) and antioxidant genes (GSTP1, NQO1, NQO2) were observed; immunohistochemistry results confirmed overexpression of GSTP1, NQO1, and DUOX2, along with higher 8-OHdG (marker of oxidant stress) in SSL and TSA (Figure 4, *B–C*). Using data from a Gene Expression Omnibus (GEO) external dataset comprising 10 tubular adenomas (TAs, also known as conventional adenomas), we found these genes related to oxidant stress were not overexpressed in conventional adenomas, indicating this feature specifically existed in premalignant lesions developed from the serrated pathway (Figure 4, *D–E*). Moreover, we found pathway enrichment of cellular response to oxidative stress was a distinct feature of BRAF-mutant CRCs, as confirmed by both The Cancer Genome Atlas (TCGA) and GEO external datasets (Figure 4, *F–G*).

We also identified upregulation of a novel tumor-promoting gene that regulated ROS level, namely SERPINB6 (encoded SerpinB6), in epithelia of SLs (Figure 4, *H* and Figure 5, *A*), whose role has not been explored in colon tumorigenesis yet. SerpinB6 was found overexpressed in Epi-SL cluster (Figure 4, *C*) and was specifically overexpressed in SSL but not in conventional adenomas (Figure 4, *D* and Figure 5, *B*). Its overexpression in SLs was confirmed by quantitative real-time polymerase chain reaction (RT-PCR) and immunohistochemistry (Figure 5, *C–D*). Of note, SerpinB6 showed significantly higher expression in BRAF-mutated CRCs in TCGA dataset (Figure 5, *E*), suggesting it might play a role in serrated tumorigenesis. Results of EdU and transwell assay in RKO cell line (a colon cancer cell line with BRAF mutation) revealed that knock down of SerpinB6 significantly attenuated the cell growth and migration ability (Figure 5, *F–H*). Furthermore, we found knock down of SerpinB6 resulted in increased reactive oxygen species (ROS) levels as indicated by 2'-7'dichlorofluorescin diacetate (DCFH-DA) assay, whereas usage of N-acetyl L-cysteine (a ROS scavenger) could partially rescue the proliferation ability in the SerpinB6-knockdown cells, suggesting that SerpinB6 can promote serrated tumorigenesis via modulating ROS levels (Figure 4, *H* and Figure 5, *I*).

Subclustering of Epi-SL Identifies One Cluster of Ki67^{hi} OLFM4^{hi} Epithelial Cells With Activation of Notch Signaling Pathway in SSL, and Another TSA-specific Cluster Producing Lysosome

To further characterize proliferation ability of epithelia in SLs, we performed cell cycle analysis (Figure 6, *A*). We found the left part cells of Epi-SL on the UMAP plot were in

G2/M phase, whereas the majority of cells in the right part were in G1 phase. Considering the high heterogeneity within the Epi-SL cluster, we generated 2 subclusters (Epi-1 and Epi-2) derived by reclustering these cells (Figure 6, *B*). Among these, thenEpi-1 subcluster was mainly found in SSLs and accounted for the majority of epithelial cells in SSLD, whereas Epi-2 was exclusively found in TSAs (Figure 6, *C*). Trajectory analysis indicated that transit amplifying cells and intermediate cells possess the ability to differentiate toward Epi-1 and Epi-2 cell population (Figure 6, *D*). Moreover, we found that the SSL-associated epithelial cluster (Epi-1) showed high expression of a set of genes related to serrated polyposis syndrome germline predisposition as reported by Soares de Lima et al.[19] (Figure 6, *E*); this Epi-1 cluster also displayed significantly higher copy number variation level (Figure 6, *F–G*), indicating its malignant potential and critical role in the formation of serrated polyp.

Compared with the Epi-2 subcluster, Epi-1 showed higher expression of proliferation marker MKI67, SOX4, stem cell marker OLFM4, HES1, and JUN, along with downregulation of LYZ, CEACAM5, DUOX2, and S100A11 (see Figure 6, *H–I*, and Supplementary Table 5).

Upregulation of the genes related to proliferative ability in the SSL-specific Epi-1 cell cluster partially explained why SSL can rapidly develop into serrated CRC. As expected, these proliferative genes showed significantly higher expression in the epithelia of SSLD than those of SSL (Supplementary Table 6). Considering its clinical importance for SSL malignant transformation, we focused on pathway alterations in the Epi-1 subcluster (MKI67^{hi} OLFM4^{hi} SSL-specific). Gene set variation analysis (GSVA) indicated several activated oncogenic and metabolic signatures in this subcluster, such as Myc, G2M checkpoint, Notch signaling, E2F targets, MAPK targets, and FOXO mediated transcription of oxidative stress metabolic genes (Figure 7, *A*). Besides, transit amplifying cells and Epi-1 cells shared commonly activated signatures of glycolysis, citrate cycle, fatty acid metabolism, arginine metabolism, and folate metabolism, suggesting metabolic reprogramming in these cells.

Based on these findings, we hypothesized that Notch signaling pathway and its downstream target MYC, which represented hallmark pathways in Epi-1, might be involved in the malignant transformation of SSL. To confirm the single-cell findings, we recalculated GSVA scores of hallmark and curated gene sets that were elevated in Epi-1, in an external bulk RNA-seq dataset comprising of SSL and SSLD samples. Strikingly, these Epi-1-elevated pathways were further enhanced in the dysplastic lesions, suggesting these features of Epi-1 were relevant to cancerous change of SSLs

Figure 1. (*See previous page*). **General clinical and single-cell transcriptional profile of SL samples included in this study.** *A*, Schematic diagram of scRNA-seq workflow in 16 SLs. *B*, Representative endoscopic image and the hematoxylin and eosin (H&E) staining of HP samples in this study. *C*, Representative endoscopic image and the H&E staining of SSL samples in this study. *D*, Endoscopic image and the H&E staining of SSL with dysplasia in this study. *E*, Representative endoscopic image and the H&E staining of TSA samples in this study. *F*, UMAP plot showed 6 color-coded global cell types of 60,568 cells analyzed. *G*, Fractions of epithelial, stromal, myeloid, T, B, and mast cells among normal, HP, SSL, SSLD, and TSA samples. *H*, Overview of sequencing cell number, cell proportion, and anatomical region of each sample.
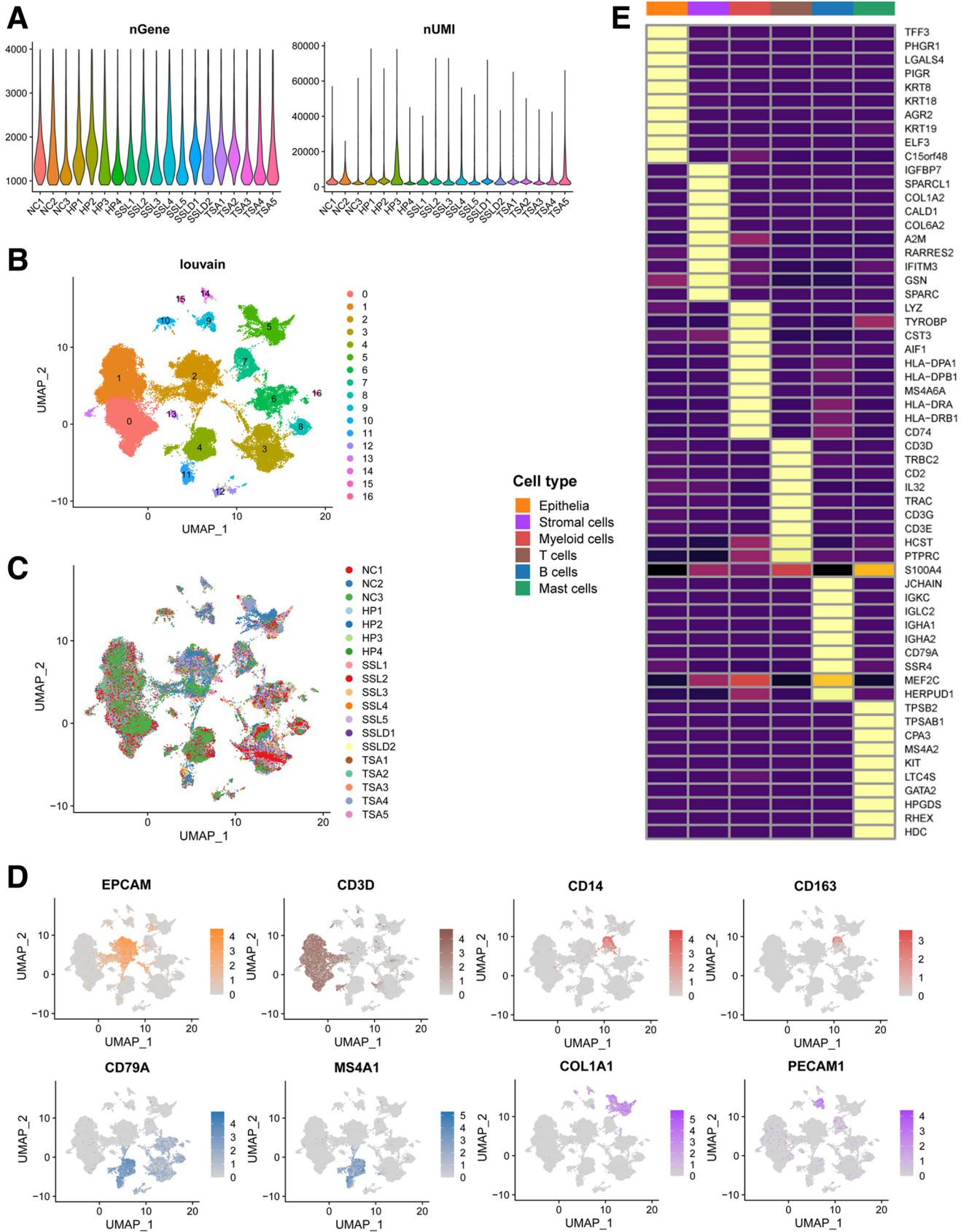
Figure 2. Dissection of the various cell types in human SLs with scRNA-seq. *A*, Violin plot showed the expressed genes (nGene) and read counts (transcript, nUMI). *B*, UMAP plot showed Louvain clusters of 60,568 cells analyzed. *C*, UMAP plot showed distribution of cells from 19 samples. *D*, Feature plots illustrated canonical markers of global cell types. *E*, Heatmap illustrated the top 10 DEGs (adjusted $P < .05$; $|\log_2 FC| > 0.25$) in each global cell type.
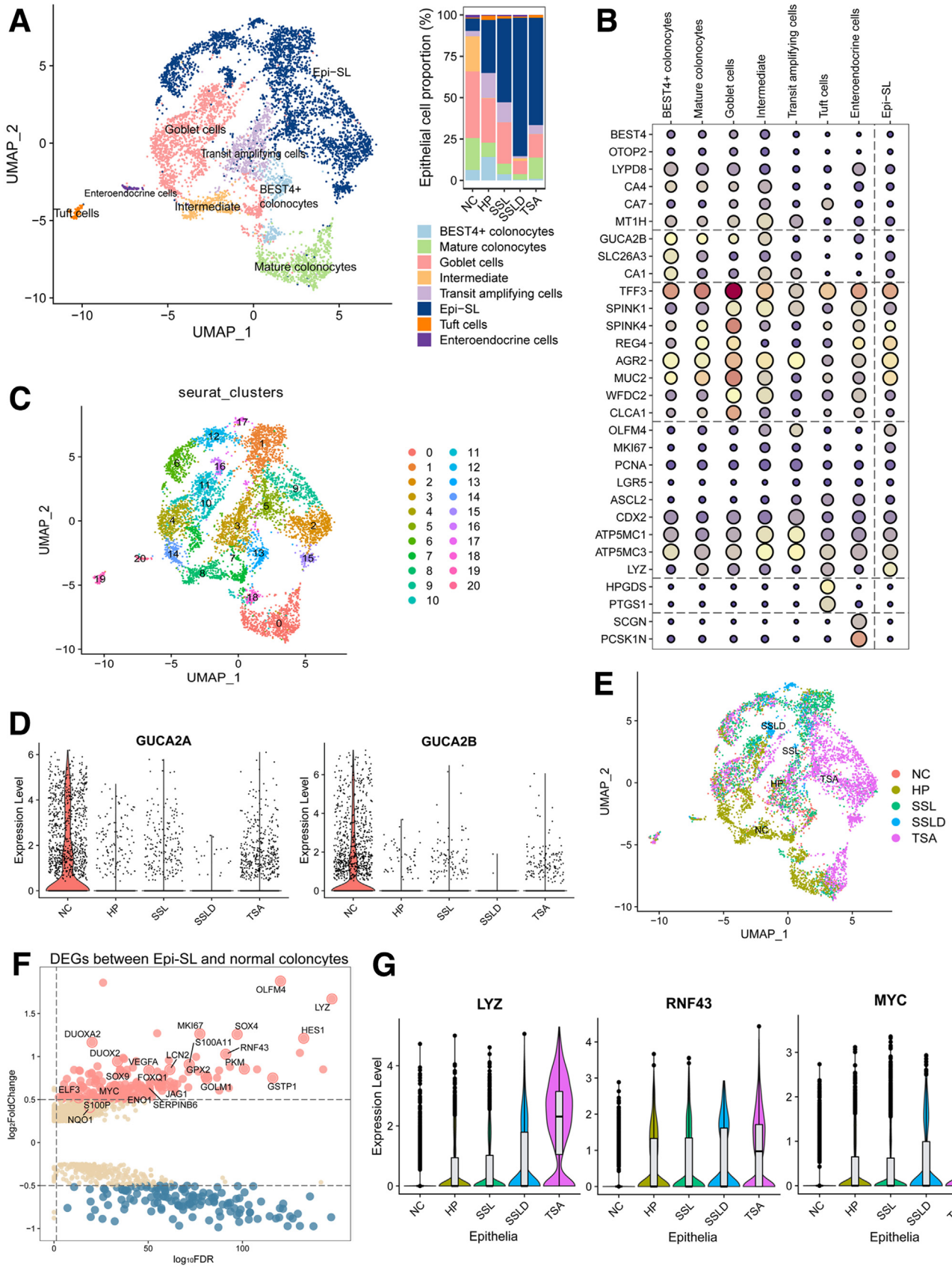
(Figure 7, *B*, top). We also identified the top 50 marker genes of Epi-1 (compared with normal colonocytes) to construct 'Epi-1 signature' (Supplementary Table 7); the GSVA score of 'Epi-1 signature' was significantly elevated in SSLs with dysplasia than those without (Figure 7, *B*, bottom). Furthermore, we performed immunohistochemistry in SSL and NC tissues. Consistent with findings of overexpression of Notch ligand, receptor, and target genes found at single-cell level in SSL epithelium (Figure 7, *C*), we observed increased expressions of JAG1 (a Notch ligand) and MYC (a Notch target) in SSL tissues at the protein level (Figure 7, *D*).

On the other hand, the TSA-associated epithelial cluster (Epi-2) was enriched in pathways of lysosome secretion, angiogenesis, and signaling by RNF43 mutants (Figure 7, *A*), which is consistent with previous reports that identified RNF43 mutation in approximately 25% of TSAs.[20] We also investigated GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment of upregulated genes in Epi-2 compared with Epi-1, and found activation of pathways related to lysosome, protein transport, neutrophil activation, calcium-dependent protein binding and S100 protein binding in Epi-2 (Figure 7, *E*). In short, Epi-2 was characterized by ectopic expression of LYZ and also higher expression of annexin and S100 family genes (Figure 7, *F* and Supplementary Table 8). Increased expression of LYZ was observed in the Epi-2 subcluster, suggesting a trend of Paneth cell differentiation in TSA. These LYZ[hi] TSA-specific epithelial cells may serve as a diagnostic biomarker for TSA. Confocal staining established the existence of abundant LYZ secretion in the crypt of TSA, but not in SSL or NC (Figure 7, *G–H*).

Subsequently, we characterized transcriptional factor (TF) regulatory network in Epi-1 and Epi-2 clusters by identifying super-regulons. As illustrated in Figure 8, upregulated TF activities for Epi-1 included IRF1, KLF10, FOSL1, RAD21, MYC, ZNF143, and BHLHE40, whereas elevated regulons for Epi-2 incorporated KLF4, KLF6, HIVEP1, NR2C2, ATF4, KLF13, and NFIL3, suggesting distinct development paths between SSL and TSA.

Both Immune Active and Immunosuppressive T Cell Components are Expanded Within Serrated Adenoma Microenvironment

Next, T cell components were subjected to clustering and UMAP visualization (Figure 9, *A*). Based on known markers, 6 main T cell types were identified (Figure 9, *B* and Supplementary Table 9). We observed a decrease in the proportion of naïve T cells in serrated neoplasms, but an increase of both CD8+ and CD4+ T cells, as well as regulatory T cells (Treg), compared with normal colon samples (Figure 9, *C*). Of note, we observed further accumulation of Tregs in SSLD, which suggested that it may facilitate carcinogenic transformation by creating an immunosuppressive microenvironment. The clinical significance of FOXP3+ Tregs was validated by immunohistochemistry analysis of expanded clinical samples of SLs (Figure 9, *D–E*), which showed a significant increase in the number of FOXP3+ cells in SSL with dysplasia samples (n = 8; *P* = .01), compared with SSLs without dysplasia (n = 8).

Because CD8+ T cells are known to play a central role in anti-tumor immunity, we focused on alterations of CD8+ T cells (including CD8+ naïve T cells, cluster 1, 6, 8, 10, 11, 12, 14, 16 in Figure 9, *A*). To investigate functional differences of CD8+ T cells in SLs and NC, we listed the DEGs between the 2 groups (Supplementary Table 10 and Figure 9, *F*). GO pathway analysis of the upregulated genes showed enrichment of pathways related to T cell activation and cytotoxic function (Figure 9, *G*). Among the DEGs, effector markers (GZMB, GZMK, GZMA, PRF1, TNF, KLRB1, BHLHE40, and NKG7) and resident markers (ITGAE, ITGB1, ITGA1, PRDM1, and CXCR6) of CD8+ T cells were significantly elevated.

Then, CD8+ T cells were reclustered to generate 5 subpopulations by clustering and dimensionality reduction from 11 Seurat clusters (Figure 9, *H–I*). Two clusters (clusters 4 and 10) expressing high levels of naïve markers were assigned as CD8+-naïve cells, 1 cluster (cluster 11) with high MKI67 was regarded as CD8+ cycling cells. The UMAP plot and heatmap of functional marker genes (Figure 9, *J*) also revealed 2 distinct subgroups of GZMK+ cells with cytotoxic function. However, one GZMK+ subgroup showed higher granzyme (GZMB, GZMK, GZMH), KLRG1, and interferon-γ (IFNG) levels, along with elevated ITGB2 expression, which was named 'GZMK+ Effector-1'; another GZMK+ subgroup was expanded in SSL with dysplasia and TSA microenvironment with much lower granzyme activity, which was titled 'GZMK+ Effector-2.' Notably, transcriptional traits of GZMK+ Effector-1 cells were similar to a cell population that was previously identified as 'ITGB2+ tissue resident memory T cells' in recent studies on human intestine.[21,22] Although these cells showed potential of strong cytotoxic function, their proportion was nearly unchanged in HP and SSL.

Another 2 subpopulations of CD8+ T cells (cluster 0, 2, 5, 7, and 8; on the right part of UMAP plot) displayed upregulation of integrin E (ITGAE) expression. Cells in cluster 5 manifested overexpressed immune checkpoint genes (PDCD1, CTLA4, TIGIT, and LAG3) with lower IFNG and granzyme B (GZMB) expression, and thus was named CD8+

Figure 3. (*See previous page*). **Epithelial cell subclusters in SLs.** *A*, UMAP representation of dimensionally reduced data following graph-based clustering with marker-based cell type assignments (*left*) and proportion of epithelial cell types in NC and SLs (*right*). *B*, Dot plot depicting expression levels of canonical colonocyte marker genes together with the percentage of cells expressing the marker. *C*, UMAP plot showing Seurat clusters of epithelial cells and association of Seurat clusters with defined epithelial cell types. *D*, Expression of GUCA2A and GUCA2B in epithelia of different groups. *E*, UMAP plot showing the sample origin of epithelial cells. *F*, Volcano plot depicting the differentially expressed genes between Epi-SL cluster and normal colonocytes (BEST4+ colonocytes, mature colonocytes, and goblet cells). *G*, Expression of LYZ, RNF43, and MYC in the epithelia of different groups.
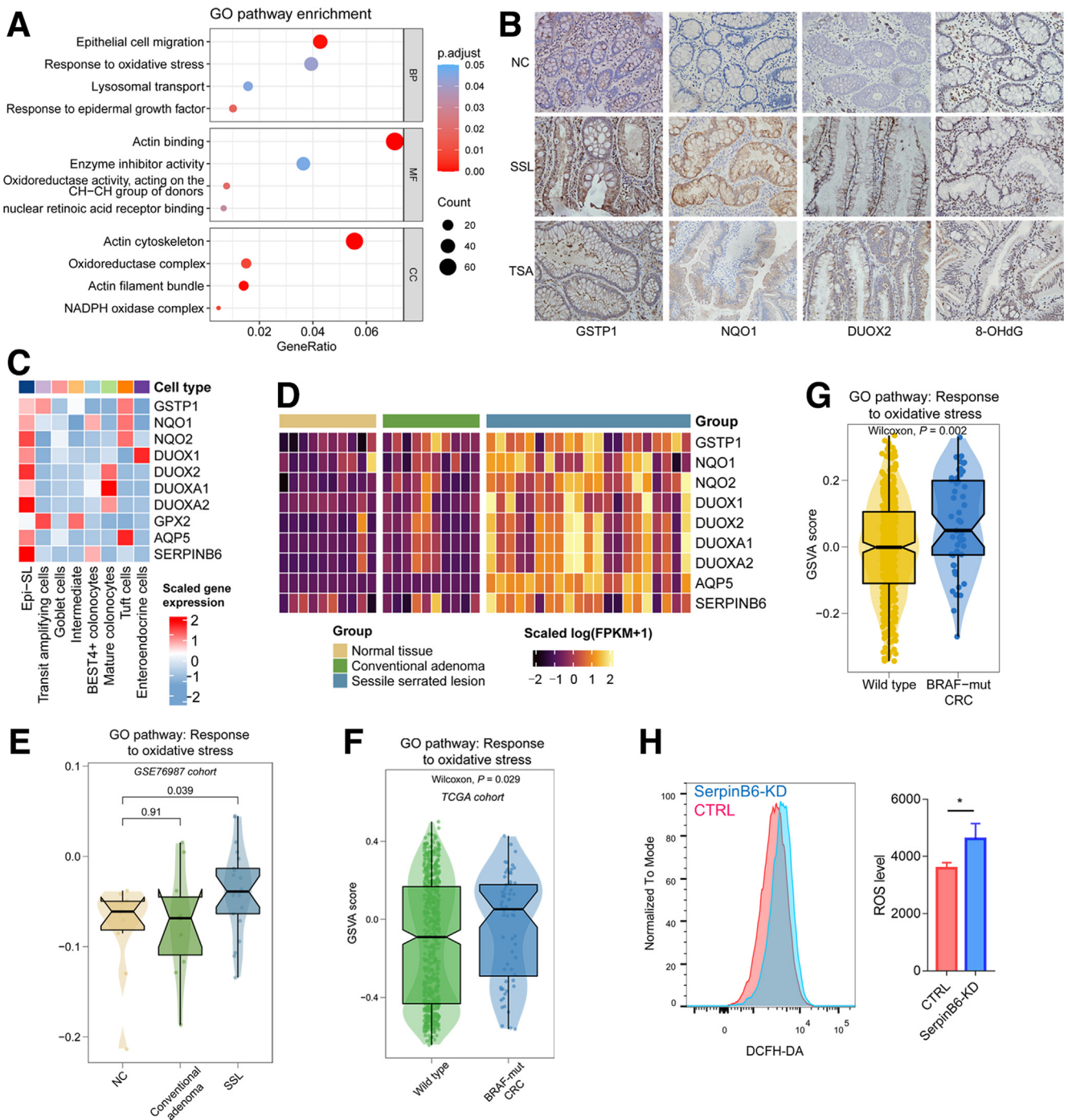
Figure 4. Identification of an Epi-SL cluster with dysregulated expression of genes related to redox balance in serrated polyps. *A*, GO enrichment analysis showed increased oxidative stress-related pathway activity in Epi-SL cluster (scored per cell by GSVA). *B*, Representative images of GSTP1, NQO1, DUOX2, 8-OHdG immunohistological staining of FFPE tissue from NC, SSL, and TSA. *C*, Heatmap showing expression of oxidative and anti-oxidative genes for each cluster. *D*, Expression of oxidative stress-related genes in NC (n = 10), conventional adenomas (n = 10), and SSLs (n = 21). Data was curated from GSE76987, which included RNA-seq data of various right-sided colon tissues. *E*, GSVA scores of response to oxidative stress pathway in NC, conventional adenomas, and SSL. Data were obtained from GSE76987. *F*, GSVA scores of response to oxidative stress pathway in BRAF wild-type (n = 634) and BRAF mutant CRCs (n = 64). Data were obtained from TCGA. *G*, GSVA scores of response to oxidative stress pathway in BRAF wild-type (n = 461) and BRAF mutant CRCs (n = 51). Data were obtained from GSE39582. *H*, SerpinB6 knockdown resulted in increased ROS levels in RKO cell line. *$P < .05$.
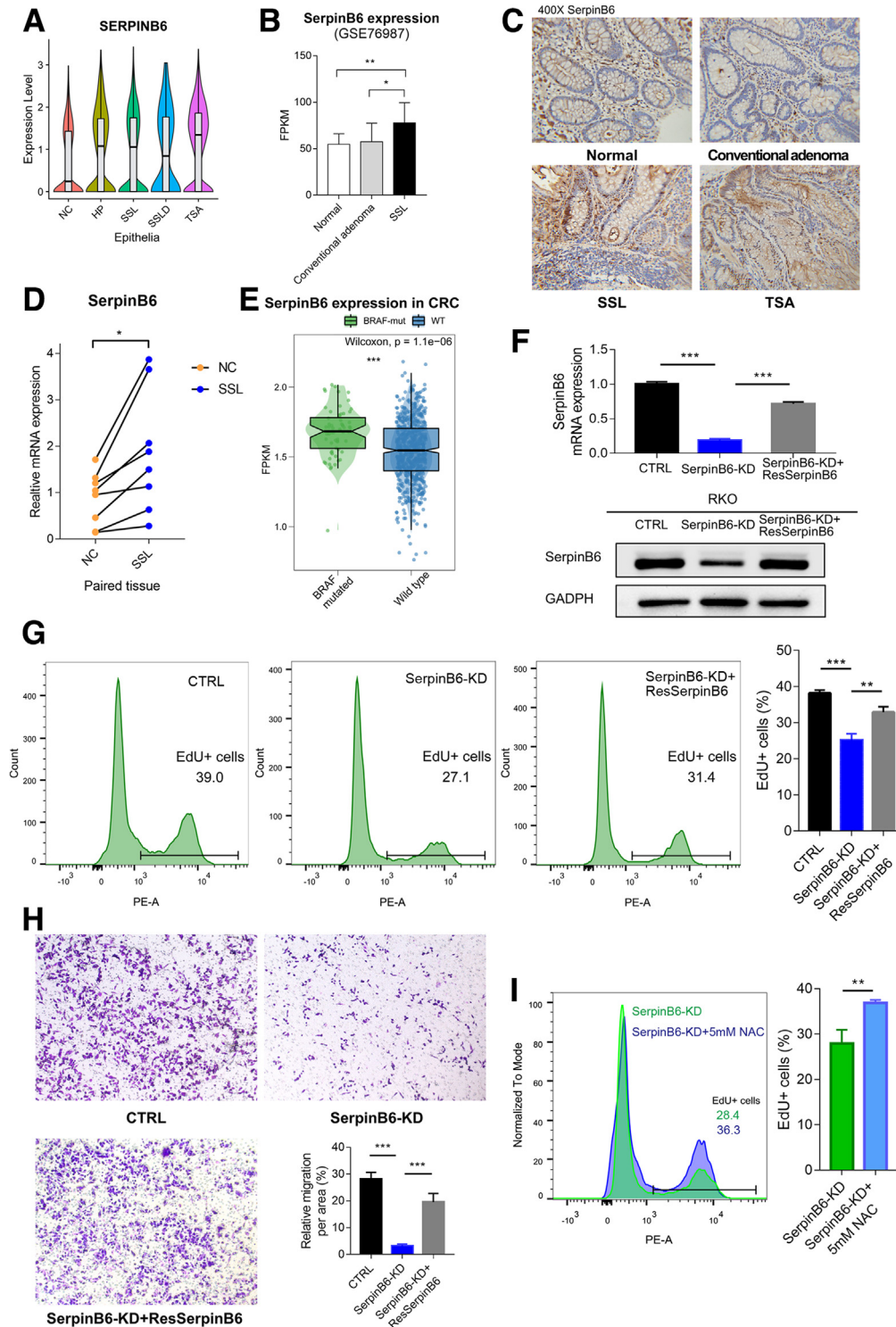
Figure 5. SerpinB6 was linked to serrated route of colon tumorigenesis and promoted colon cancer cell line proliferation. *A*, Violin plot showed increased expression level of SerpinB6 in epithelia of HP, SSL, SSLD, and TSA, compared with normal tissue. *B*, Upregulation of SerpinB6 in SSL (n = 21) compared with tubular adenoma (conventional adenoma, n = 10) in a bulk RNA-seq dataset (GSE76987). *C*, Immunohistochemistry staining of SerpinB6 in normal, conventional adenoma, and SLs. *D*, Quantitative RT-PCR analysis of SerpinB6 expression in 8 fresh SSL tissues compared with paired normal tissues. *E*, SerpinB6 expression in BRAF-mutant (n = 64) and wild-type (n = 634) CRCs in TCGA database. *F*, Knock down (SerpinB6-KD) and re-introduction (ResSerpinB6) of SerpinB6 in RKO cell line, as validated by mRNA (n = 3 replicates/group for quantitative RT-PCR assay) and protein level. *G*, EdU assay showed knock down of SerpinB6 suppressed cell proliferation of RKO, while re-introduction of SerpinB6 rescued cell growth (n = 3 replicates/group). *H*, Transwell experiment showed that knock down of SerpinB6 suppressed migration ability of RKO, while re-introduction of SerpinB6 rescued cell migration (n = 3 replicates/group). *I*, Treatment of 5 mM N-acetyl L-cysteine (NAC) overnight partially rescued the proliferation ability of RKO after knocking down SerpinB6 (n = 3 replicates/group). *$P < .05$; **$P < .01$; ***$P < .001$.
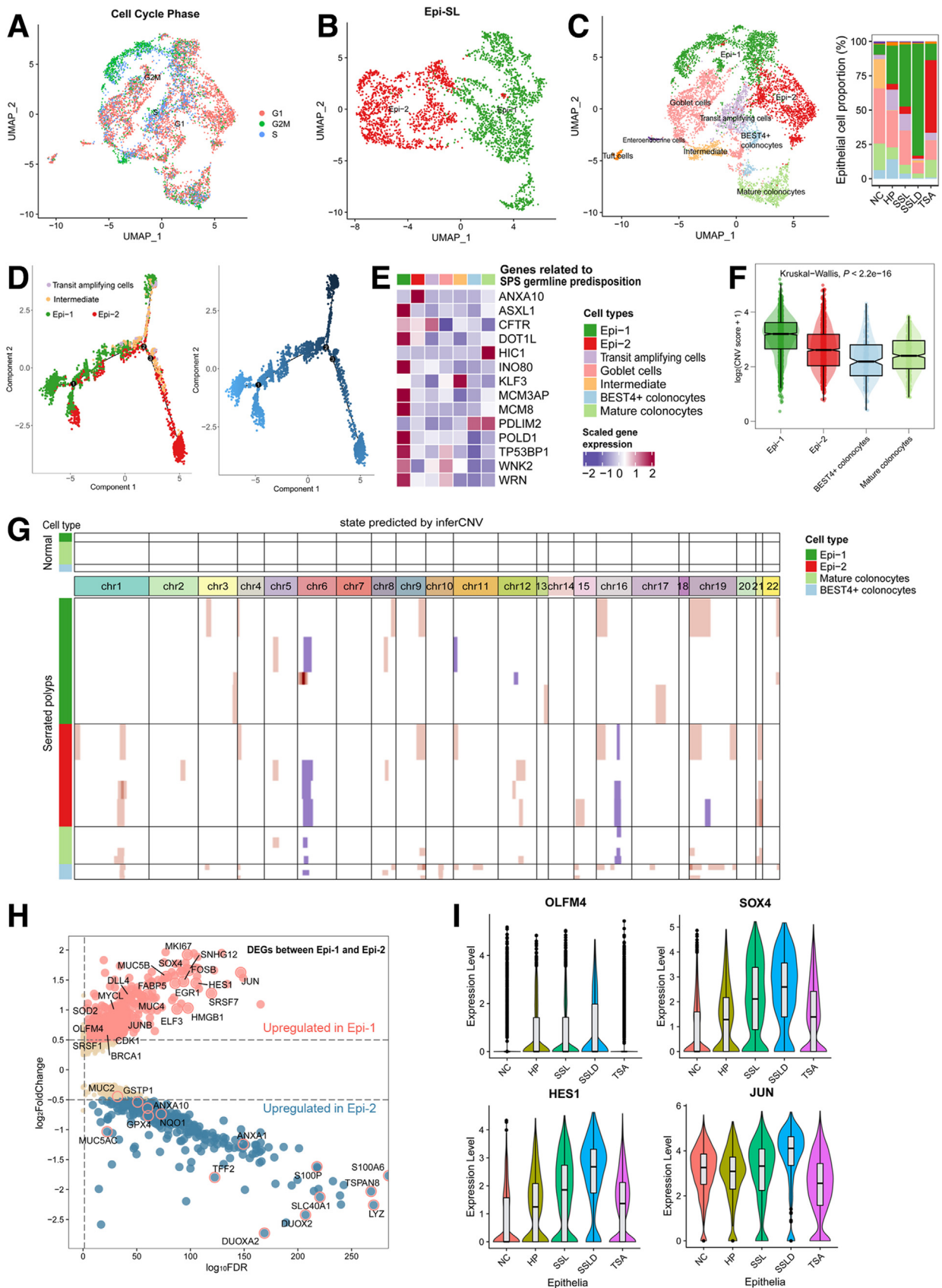
Exhausted cluster, which was mainly detected in SSL. Interestingly, the other cell component (cluster 0, 2, 7, and 8) with high expression of ITGAE (encoded CD103), which serves as a marker gene of tissue-resident memory T cells (Trm) with limited recirculation,[23] was significantly enriched in the microenvironment of HP, SSL, SSL with dysplasia, and TSA (Figure 10, A); we designated these cells as CD103$^+$ Trm cells. These CD103$^+$ Trm cells showed high expressions of IFNG, GZMB, and IL2. Consistent with these findings, using a predefined Trm CD8$^+$ gene signature by Savas et al,[23] we demonstrated that GSVA scores of Trm CD8$^+$ gene signature were significantly elevated in CD8$^+$ T cells in all 4 kinds of SLs (Figure 10, B) (all $P < .001$). Cell cycle analysis revealed that G2/M proportion of CD103$^+$ Trm was not increased in SLs, whereas G2/M proportion of naïve CD8$^+$ T was augmented, suggesting that the increase of these cells was probably attributable to differentiation from naïve cells (Figure 10, C).

Trm Metabolic Phenotypes Suggest Retinoic Acid as a Modulator of Antitumor Immune Response in Serrated Tumorigenesis

Trm cells play a central role in anti-tumor immunity with enhanced cytotoxic function and favorable prognosis.[23,24] In our study, we observed marked enrichment of CD103$^+$ Trm component in SLs; therefore, we further explored the clinical importance and potential regulating factors of these cells. We confirmed co-localization of CD8 and CD103 by confocal immunofluorescence microscopy in SSL slides (Figure 10, D). Next, we evaluated CD103 expression in NC, serrated tumors, and TA (conventional adenoma) by immunohistochemistry (Figure 10, E). In serrated tumors, CD103 was exclusively expressed in lymphocytes; we observed significantly increased number of CD103$^+$ cells in HP ($P = .01$), SSL ($P = .002$), SSLD ($P = .003$), TSA ($P = .01$), and BRAF-mutant CRC ($P = .01$), compared with NC (Figure 10, F). In contrast, CD103$^+$ cells were not augmented in conventional adenoma ($P = .40$).

Considering that previous reports showed that the number of CD8$^+$CD103$^+$ T cells decreases from ascending colon to rectum[25,26] and NC location could cause confounding bias, we additionally collected biopsies of colorectal healthy mucosa from ascending colon, transverse colon, descending colon, and rectum (n = 6 for each group),

and performed the immunohistochemical staining of CD103 (ITGAE). The results showed that the number of CD103$^+$ cells decreased from ascending colon to rectum ($P < .001$) (Figure 11, A–B). However, numbers of CD103$^+$ cells in HPs (45.1 ± 15.7), SSLs (58.0 ± 19.7) and TSAs (42.9 ± 17.1) were still higher, even compared with normal mucosa from ascending colon (21.6 ± 2.2). Similar results were observed in the Human Protein Atlas database, which showed that CD103$^+$ cells decreased significantly from small intestine to large intestine (Figure 11, C). However, the difference between colon and rectum was less striking. That is to say, the fluctuation of T cells in different locations of the large intestine do not contribute to our main findings of abundant CD103$^+$ Trm cells in serrated polyps.

Then, we determined the immunological features of colon CD8$^+$ Trm cells by flow cytometry (FCM) in another 8 fresh SSL tissues. In activated CD8$^+$CD69$^+$ T cells of SSL, we found that CD103$^+$ Trm expressed higher levels of GZMB compared with CD103$^-$ cells, whereas ITGB2$^+$ Trm (GZMK$^+$ Effector-1 cluster in our study) expressed higher levels of GZMK, GZMB, and KLRG1 (Figure 11, D–E), consistent with the scRNA-seq data. These results suggested that CD103$^+$ and ITGB2$^+$ Trm cells have higher cytotoxic potential compared with non-Trm cells in the CD8$^+$ T cell responses of SLs.

Further, using 2 external cohorts, we found the Trm CD8$^+$ gene signature[23] was significantly enriched in BRAF-mutant CRCs in TCGA (both $P < .001$) (Figure 11, F–G), suggesting its close association with serrated pathway. Subgroup analysis stratified by BRAF and MSI status showed higher Trm signature score in MSI-H BRAF-mutant (n = 71), compared with MSS BRAF-mutant CRCs (n = 33; $P = .004$), as well as a trend of higher Trm signature score in MSS BRAF-mutant, compared with MSS BRAF-wildtype CRCs ($P = .08$) (Figure 11, H). We also investigated the prognostic significance of these CD8$^+$ T cell populations in 263 patients with CRC with BRAF or KRAS mutation, using the signature genes for each CD8$^+$ T cell type identified from scRNA-seq (Supplementary Table 11). The CD103$^+$ Trm and CD8$^+$ cycling clusters were associated with favorable outcomes ($P = .005$ and $P = .001$, respectively), whereas GZMK$^+$ Effector-1 showed potential protective effect ($P = .058$) with marginal significance (Figure 11, I).

Immunometabolism was shown to be a promising target for improving anti-tumor immunity.[27] To understand the metabolic reprogramming in the tumor microenvironment

---

Figure 6. (*See previous page*). **Identification of two distinct epithelial cell populations (Epi-1 and Epi-2) in SSL and TSA.** *A*, Cell cycle analysis of epithelia in SLs. *B*, Reclustering of Epi-SL resulted in 2 subclusters. *C*, Proportion of Epi-1 and Epi-2 subclusters among HP, SSL, and TSA. *D*, Unsupervised trajectory plot inferred by Monocle 2 showed that Epi-1 and Epi-2 cells derived from transit amplifying and intermediate cells. *E*, Higher expression of genes related to SPS germline predisposition in Epi-1 subcluster. *F*, CNV score calculated by inferCNV R package among four different cell types (BEST4$^+$, mature colonocytes, Epi-1, and Epi-2) in serrated polyps. *G*, Chromosomal landscape of inferred large-scale CNVs distinguishing serrated polyp epithelial cells from normal colonic cells. The SL epithelial cells were shown with different cell types (BEST4$^+$, mature colonocytes, Epi-1, and Epi-2) and chromosomal regions (horizontal). Amplifications (*red*) or deletions (*blue*) were plotted, with normal sample as the reference (*top*). A 6-state CNV model that predicts the following CNV levels was used: 0: complete loss; 0.5: loss of 1 copy; 1: neutral; 1.5: addition of 1 copy; 2: addition of 2 copies; 3: addition of more than 2 copies. *H*, Volcano plot of DEGs between Epi-1 and Epi-2 (adjusted $P < .05$; |log$_2$FC| > 0.25). *Red points* represented genes significantly overexpressed in Epi-1 cluster, and *blue points* showed genes significantly overexpressed in Epi-2 cluster. *I*, Violin plot showed expression of OLFM4, SOX4, HES1, and JUN in epithelia of NC, HP, SSL, SSLD, and TSA.
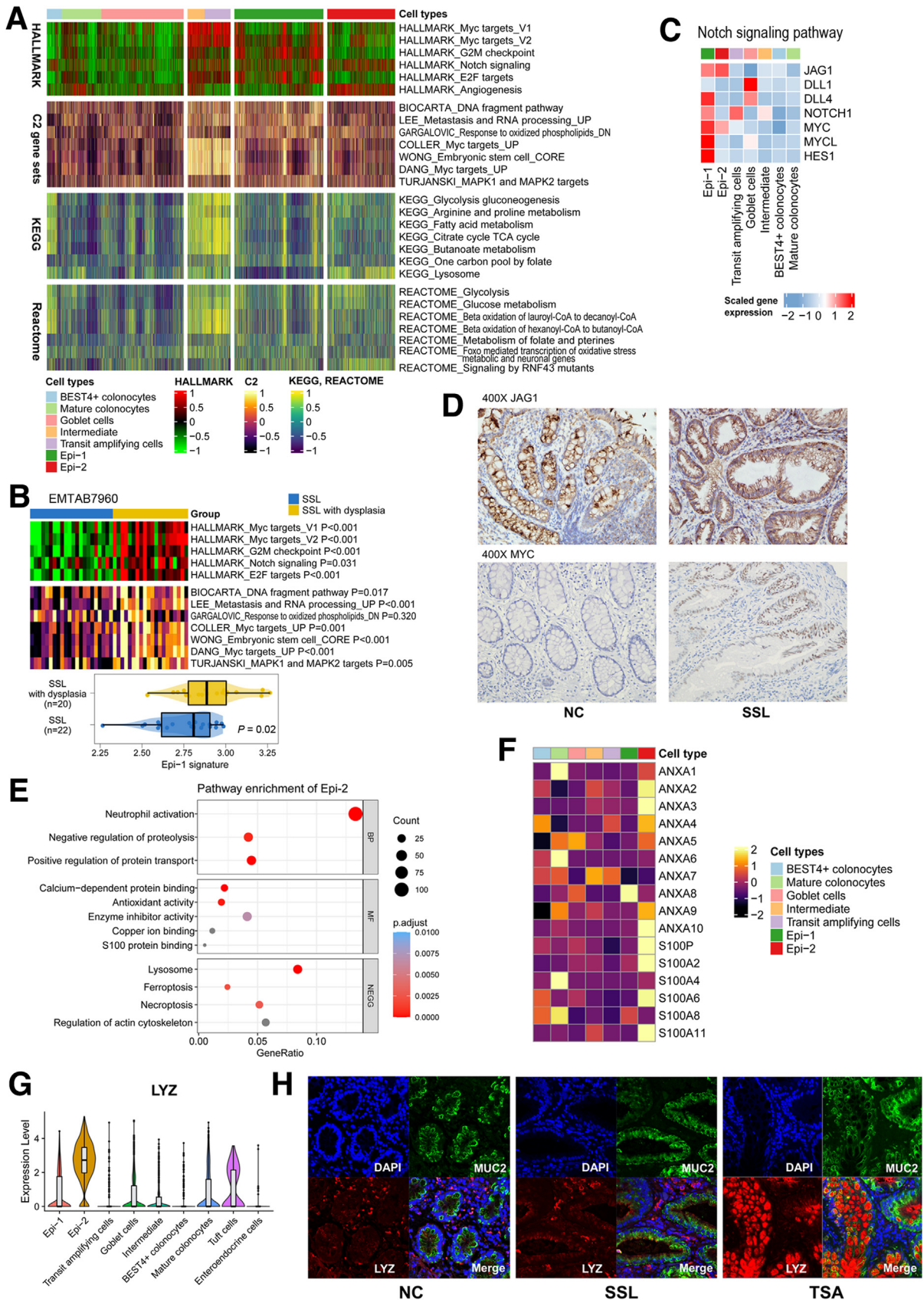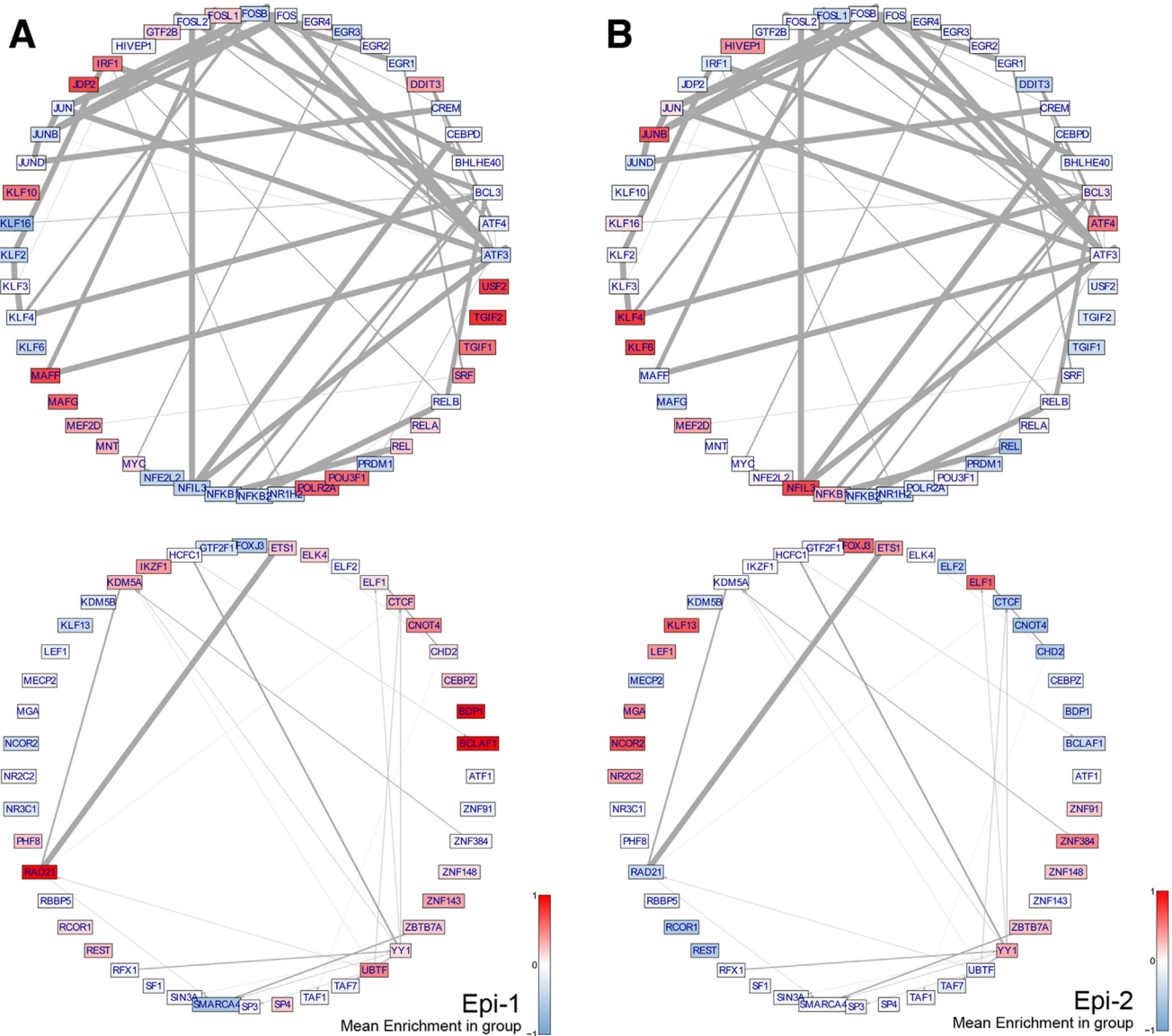
Figure 8. TF target network created from normal and SL cells, organized into super-regulons for Epi-1 (*A*) and Epi-2 (*B*) cells. Color in each vertex represented enrichment scores in corresponding regulon, whereas edge opacities were determined by the inferred TF-target weightings between regulons.

(TME), we employed the scMetabolism tool[28] for quantifying single-cell metabolism of CD8$^+$ T cells. As shown in Figure 11, *J*, we identified remarkable activation of retinol metabolism in CD103$^+$ Trm cells, but this phenomenon was not observed in the other activated CD8$^+$ T cells. To test the effect of retinol metabolism on differentiation from naïve T cells to CD103$^+$ Trm, we added retinoic acid (RA), the product of retinol metabolism, in a well-established in vitro Trm differentiation system (Figure 12, *A*; see Materials and Methods section). We found that addition of both 10 nM and 100 nM RA significantly induced more CD69$^+$CD103$^+$ T cells (Figure 12, *B–D*). We also observed significant

---

Figure 7. (*See previous page*). **Distinct subtypes of epithelial cells in SSL and TSA tissues.** *A*, Comparison of pathway activities (calculated based on GSVA) among different epithelial cell subtypes. The scores of pathways were normalized. Each column represents a single cell and each cell cluster analyzed was color-coded. *B*, Higher enrichment of pathways upregulated in Epi-1 was also observed in SSL with cytological dysplasia samples (n = 22, data from EMTAB7960), compared with SSL without dysplasia (n = 20; top). Higher Epi-1 signature score was observed in SSL with cytological dysplasia samples (bottom). *C*, Heatmap showed expression of genes related to Notch signaling pathway in each epithelial cell type. *D*, Immunohistochemistry staining of JAG1 and MYC in NC and SSL tissue. *E*, GO pathway enrichment analysis of upregulated genes in Epi-2. *F*, Heatmap showing expression of annexin and S100 family genes in epithelial subclusters. *G*, Violin plot showed the highest expression of LYZ in Epi-2 cluster. *H*, Confocal staining of MUC2 and LYZ in NC, SSL, and TSA tissue.

upregulation of SDR16C5 (a rate-limiting retinol dehydrogenase in the biosynthesis of RA) in the epithelia of SLs but not in conventional adenomas (Figure 12, *E–G*), as well as in

BRAF-mutant CRCs (Figure 12, *H*), suggesting a potential role of epithelium-derived RA in the modulation of the TME in SLs.
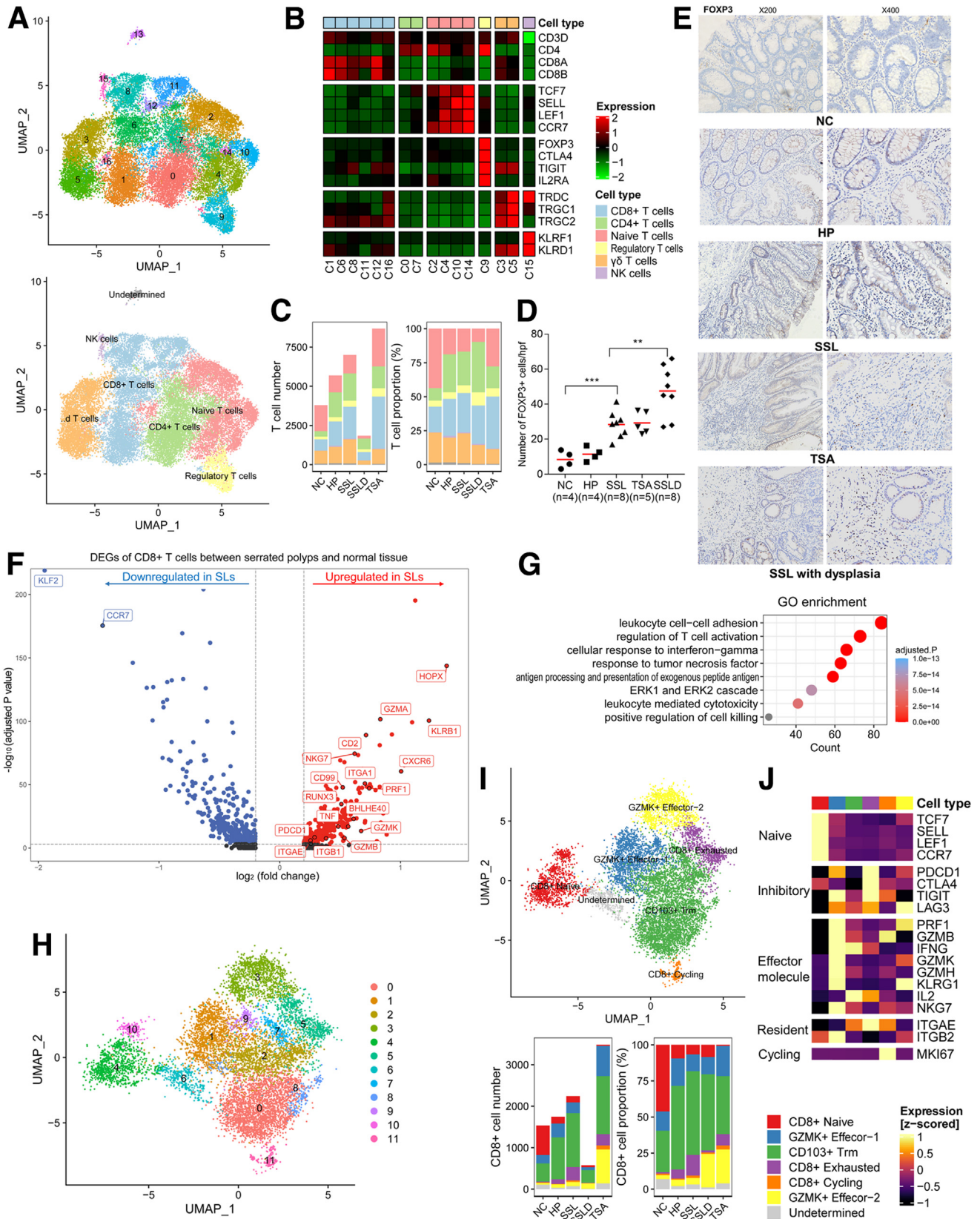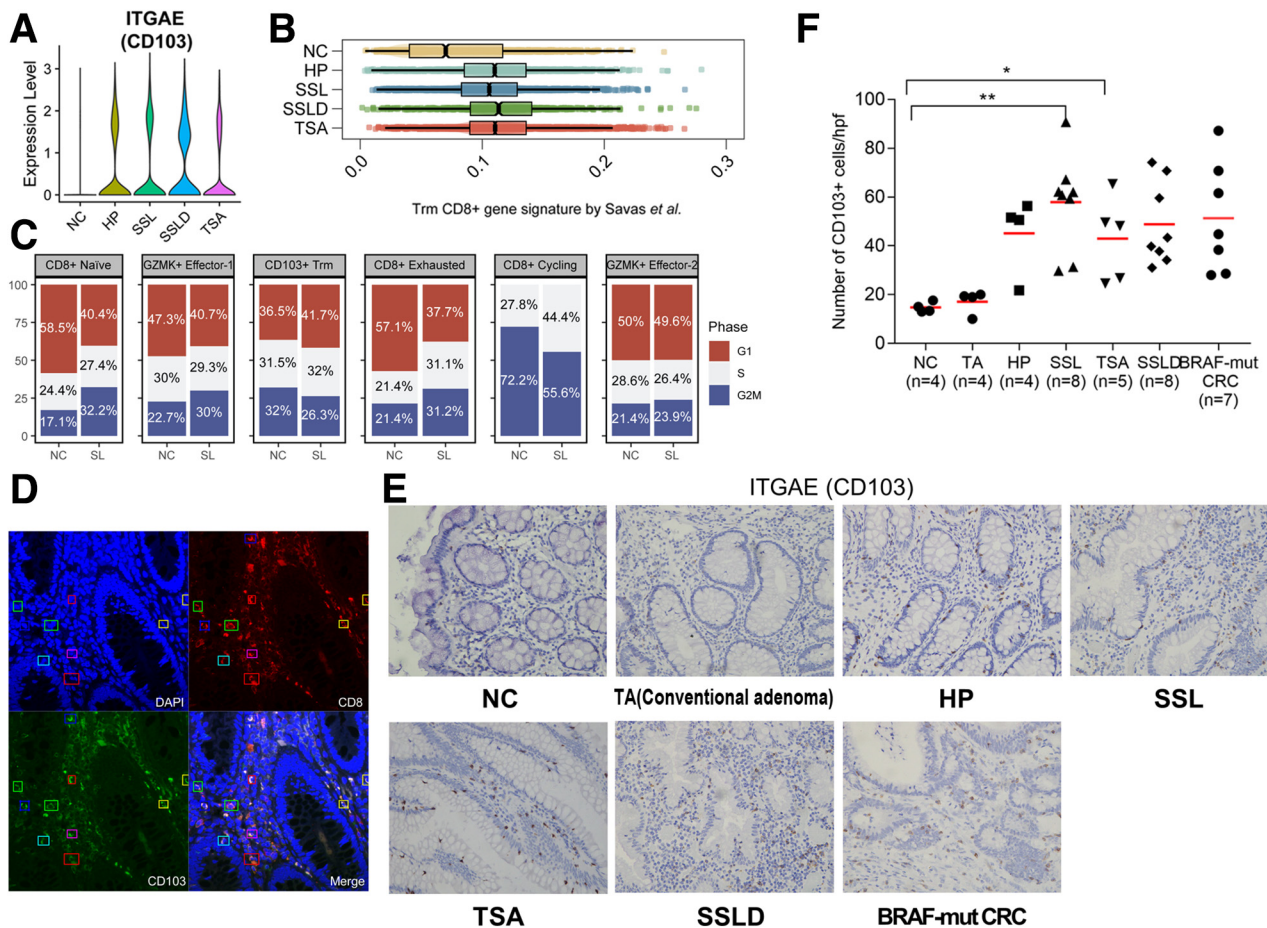
Figure 10. Characteristics of CD103$^+$ Trm cells in SLs. *A*, Violin plot showing the expression of ITGAE/CD103 in CD8$^+$ T cells in NC and SLs. *B*, GSVA scores of Trm CD8$^+$ gene signature in NC and SLs. *C*, Cell cycle analysis of different CD8$^+$ T cell subclusters in NC and SLs. The full list of DEGs is summarized in Supplementary Table 10. *D*, Confocal staining of CD103 (*in green*) and CD8 (*in red*) in SSL. *E*, Representative immunohistochemical staining of CD103 in NC, SLs, conventional adenomas (TAs), and BRAF-mutant CRCs. *F*, Quantitative of CD103$^+$ cells in NC, SLs, conventional adenomas, and BRAF-mutant CRCs. *$P < .05$; **$P < .01$.

## Altered Myeloid and B Cell Composition Across the SL Spectrum

We identified 2 macrophage clusters and 2 dendritic cell (DC) clusters in SLs (Figure 13, *A–C*). Among 2 CD14$^+$ macrophage subgroups, FCN1$^+$ inflammatory macrophages expressed various proinflammatory genes, including FCN1, VCAN, S100A8, S100A9, and IL1B, whereas another subgroup of anti-inflammatory macrophages expressed typical M2 markers such as MRC1, CD163, C1QA, APOE, SELENOP, and MAF. The 2 DC clusters, namely conventional and plasmacytoid DC, were designated by high expression levels of CD1C, CLEC9A, FCGR1A, and MZB1, respectively. GSVA analysis demonstrated that inflammatory response-related pathways were enriched in FCN1$^+$ inflammatory macrophages, whereas these were downregulated in the anti-inflammatory macrophage group. The 2 DC clusters

Figure 9. (*See previous page*). **Transcriptomic heterogeneity of T cells in the tumor microenvironment of SLs.** *A*, UMAP plot for clustering of T cells, color-coded for 6 clusters. *B*, Heatmap showed expression of marker genes in each T cell cluster: CD8$^+$ T cells (cluster 1, 6, 8, 11, 12, 16), CD4$^+$ T cells (cluster 0, 7), naïve T cells (cluster 2, 4, 10, 14, with high expression of naïve markers TCF7, SELL, LEF1, CCR7), Treg (cluster 9, high expression of FOXP3), $\gamma\delta$ T cells (cluster 3, 5, high expression of TRDC1 and TRGC1), and NK cells (cluster 15, high expression of KLRF1). *C*, Cell number and proportion of 6 T cell clusters among HP, SSL, and TSA. *D*, Comparison of numbers of FOXP3$^+$ cells in NC, HP, SSL, TSA, and SSLD tissues. *E*, Immunohistochemistry staining of regulatory T cells (FOXP3$^+$ cells) in NC and SLs. *F*, Volcano plot displayed DEGs of CD8$^+$ T cells between SLs and normal tissue. *G*, GO pathway enrichment analysis of upregulated genes in CD8$^+$ T cells of SLs. *H*, UMAP plot showing Seurat clusters of CD8$^+$ T cells. *I*, UMAP plot for subclustering of CD8$^+$ T cells, color-coded for 6 subclusters (*top*). Cell number and proportion of CD8$^+$ T cell subclusters among HP, SSL, and TSA (*bottom*). *J*, Heatmap showed expression of marker genes related to naïve, inhibitory, effector, resident, and cycling function in each CD8$^+$ T cell subcluster. **$P < .01$; ***$P < .001$.

showed enrichment of pathways related to antigen presentation and hypoxia (Figure 13, *D–E* and Supplementary Table 12).

Interestingly, expansion of M2-like anti-inflammatory macrophages was observed in SLs, particularly in SSL with dysplasia and TSA. Moreover, the proportion of cDCs
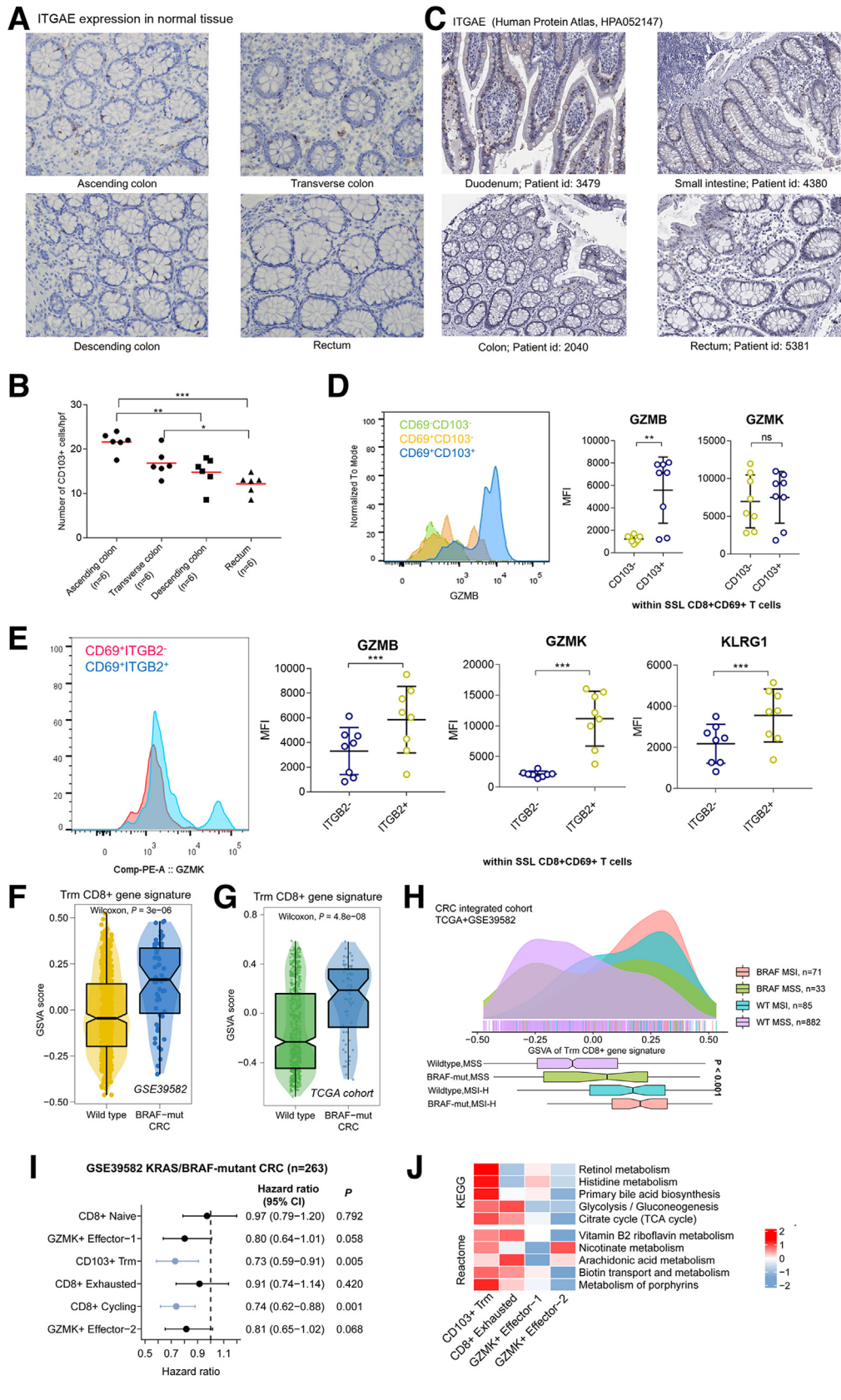
was increased in SSL, but decreased in TSA. To further investigate the TFs regulating this myeloid differentiation process, we performed single-cell regulatory network inference and clustering (SCENIC) analysis (Figure 13, F). Several known and novel TFs regulating anti-inflammatory macrophage (M2) differentiation were identified. In accordance with previous reports,[29,30] MAF, SPIC, NR1H3 (LXRα) were identified as candidate TFs for anti-inflammatory macrophages. We also identified some TFs whose roles in macrophage differentiation are not known, such as ETV5 (Figure 13, G). These results provided potential targets for inhibiting or reversing the aberrant polarization of immunosuppressive macrophages in serrated TME.

As shown in Figure 14, A–B, we clustered B cell population (CD79A$^+$) into plasma cells (SDC1$^+$) and non-plasma B cells (CD19$^+$MS4A1$^+$). In large intestine, we found that the vast majority of plasma cells showed high expression of IGHA1, and thus we named this cluster IgA$^+$ plasma cells (Figure 14, C and Supplementary Table 13). In SSL and TSA, we found decreased proportion of non-plasma B cells, but higher proportion of IgA$^+$ plasma cells (Figure 14, B). Besides the production of antibodies, tumor-infiltrating B cells also produce cytokines and facilitate antigen presentation and formation of tertiary lymphoid structures via CXCL13-CXCR5 axis.[31] Compared with non-plasma B cells in normal tissue, naïve mature B cells and memory B cells in SSL and TSA displayed diminished expressions of TNF-α and CXCR5, implying weakened antitumor secretion and recruitment ability of tumor-infiltrating B cells to form tertiary lymphoid structures. We also found higher expression of Absent in melanoma 2 (AIM2), an inflammasome component, in B cells within serrated adenoma. Because a recent study has uncovered the role of AIM2 regulating B-cell differentiation toward antibody-producing plasma cell via PRDM1 and XBP1,[32] upregulation of AIM2 may contribute to increased proportion of plasma cells in SLs.

To investigate altered cell-cell communications in SLs, we inferred the cell-cell interaction between plasma cells and CD8$^+$T/myeloid cells in normal tissue and serrated adenomas, respectively, using Cellchat. As shown in Figure 14, D, putative gain of midkine (MDK)-related interactions was observed between IgA$^+$ plasma cells and other immune cells (particularly myeloid cells) in SLs. MDK was reported to promote cell growth and proliferation and promote immunosuppressive macrophage differentiation;[33] we further validated the increased MDK expression on IgA$^+$ plasma cells in SSL by confocal staining (Figure 14, E). A recent study also reported a cluster of IgG4$^+$ B cell clones that produce MDK and VEGFA to promote angiogenesis.[34] The cytokine production function of B cell clusters in serrated adenoma and adenocarcinoma warrant further investigation.

To identify underlying regulators for increased IgA$^+$ plasma cells in serrated adenoma, SCENIC analysis among 4 types of B cells was performed (Figure 14, F). For IgA$^+$ plasma cell, several well-established TFs were identified (eg, CREB3, XBP1, CREB3L2, and PRDM1). Of note, ATF5 was also found to be an underlying TF regulating plasma cell differentiation, which has not been reported before. Both ATF5 and 65 ATF5 targets were predominantly expressed in IgA$^+$ plasma cell, as depicted in the UMAP plot (Figure 14, G). To validate the SCENIC results, we isolated B cells and plasma cells from the SSL tissue samples by FCM (see gating strategy in Figure 14, H) and found significantly higher median fluorescence intensity of ATF5 in plasma cells ($P = .0009$) (Figure 14, I).

PDGFRA$^+$ Fibroblasts Accumulate in Precancerous SLs and BRAF-mutant CRCs

We detected 15 sub-cell types from 5139 stromal cells (Figure 15, A–C), including 5 fibroblast types, 4 endothelial cell (EC) types, and 4 minor stromal cell types (enteric glial cells, pericytes, vascular smooth muscle cells, and endothelial progenitor cells) with distinct marker genes (Figure 15, D and Supplementary Table 14). Fibroblasts (COL3A1$^+$) in SLs were clustered into 5 subpopulations: PDGFRA$^+$ fibroblast, myofibroblast (TAGLN, ACTA2, and MYL9), Stromal-1 (high level of SFRP2, a canonical Wnt inhibitor), Stromal-2, and Stromal-3 (expressed markers of ADAMDEC1 and APOE, suggesting activated EGF signaling pathway; Stromal-3 also manifested high CCL4 and CCL5 expression). Four types of EC (ENG$^+$ and PECAM1$^+$) were assigned: tip-like EC, stalk-like EC, EC-1, EC-2. EC-1 existed predominantly in NC with upregulation of ICAM1, whereas EC-2 was mainly observed in SLs. Stalk-like EC showed lower expression of CD36, a key modulator of fatty acid metabolism.

Figure 11. CD103$^+$ Trm in SLs showed cytotoxic activity and might be regulated by retinol metabolism. *A*, Immunohistochemistry staining of CD103$^+$ cells in colorectal healthy mucosa from ascending colon, transverse colon, descending colon, and rectum. *B*, Comparison of numbers of CD103$^+$ cells from proximal colon to distal colon tissues. *C*, Expression of ITGAE/CD103 in histological sections of duodenum, small intestine, colon, and rectum, from Human Protein Atlas (HPA) database (https://proteinatlas.org). *D*, FCM analysis in activated CD8$^+$ T cells showed increased GZMB in CD103$^+$ cells, and ITGB2$^+$ cells expressed higher levels of GZMK and GZMB. *E*, FCM analysis showed increased GZMK, GZMB, and KLRG1 in ITGB2$^+$ CD69$^+$ CD8$^+$ T cells (GZMK$^+$ Effector-1 cells) within SSL tissues. *F*, GSVA scores of Trm CD8$^+$ gene signature in BRAF wild-type (n = 461) and BRAF mutant CRCs (n = 51). Data were obtained from GSE39582. *G*, GSVA scores of Trm CD8$^+$ gene signature in BRAF wild-type (n = 634) and BRAF mutant CRCs (n = 64). Data were obtained from TCGA. *H*, Distribution of GSVA scores of Trm CD8$^+$ gene signature in MSI-H BRAF-mutant (n = 71), MSS BRAF-mutant (n = 33), MSI-H BRAF-wildtype (n = 85), MSI-H BRAF-wildtype CRCs (n = 882). Data were integrated from both GSE39582 and TCGA, with batch effect corrected. *I*, The correlation between estimated CD8$^+$ T subpopulations and recurrence-free survival in 263 patients with BRAF/KRAS-mutant CRC. The hazard ratio and the *P* value was evaluated by the Cox proportional hazard model with 95% confidence interval (CI). *J*, Metabolic activities at the single-cell resolution estimated by scMetabolism in each CD8$^+$ T cell subpopulation. *P < .05; **P < .01; ***P < .001.
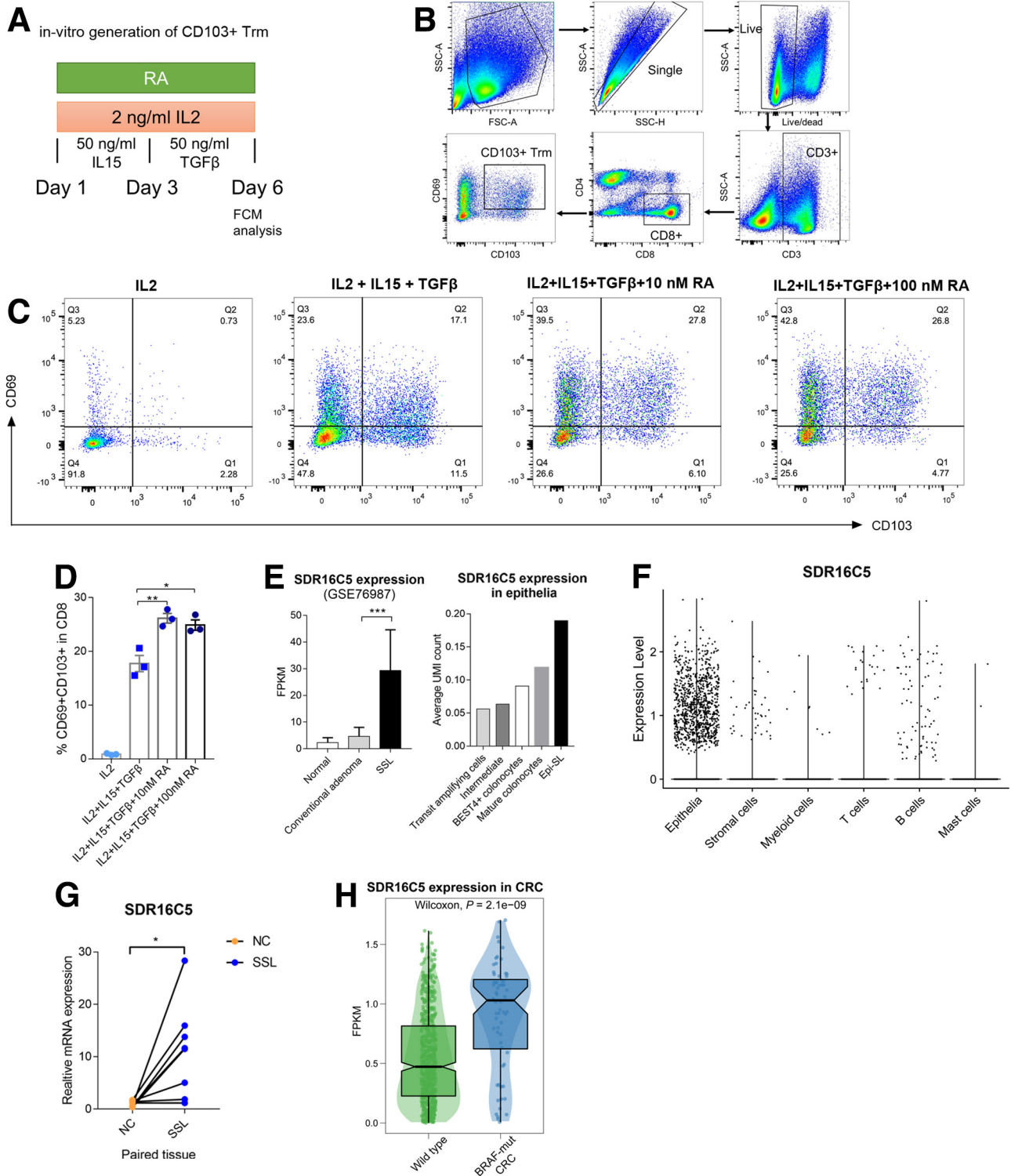
Figure 12. Effects of RA on in-vitro generation of human CD8$^+$CD69$^+$CD103$^+$ T cells. *A*, Diagram of in-vitro Trm generation assay, with addition of 10 nM or 100 nM RA. *B*, The gating strategy used for flow cytometry analysis of CD103$^+$ Trm cells. *C*, Effects of RA on in-vitro generation of human CD8$^+$CD69$^+$CD103$^+$ T cells. FCM analysis of sequential exposure to IL-15 and TGF-β induced CD8$^+$ Trm cells was performed for 3 times. RA (10 nM or 100 nM) was added to the culture medium; expansion of CD8$^+$CD69$^+$CD103$^+$ T cells was determined on day 6. A representative scatter plot of an experiment was shown (gating on CD8$^+$ T cells). *D*, Statistical analysis of the expansion of CD8$^+$CD69$^+$CD103$^+$ T cells in each group. *E*, Increased SDR16C5 expression in SSL compared with conventional adenomas (*left*), as well as in Epi-SL compared with other epithelial cell types (*right*). *F*, Expression of SDR16C5 in different global cell types showed it mainly expressed in the epithelia. *G*, Quantitative RT-PCR analysis of SDR16C5 expression in 7 fresh SSL tissues compared with paired normal tissues. *H*, SDR16C5 expression in BRAF-mutant (n = 64) and wild-type (n = 634) CRCs in TCGA database. *$P < .05$; **$P < .01$; ***$P < .001$.
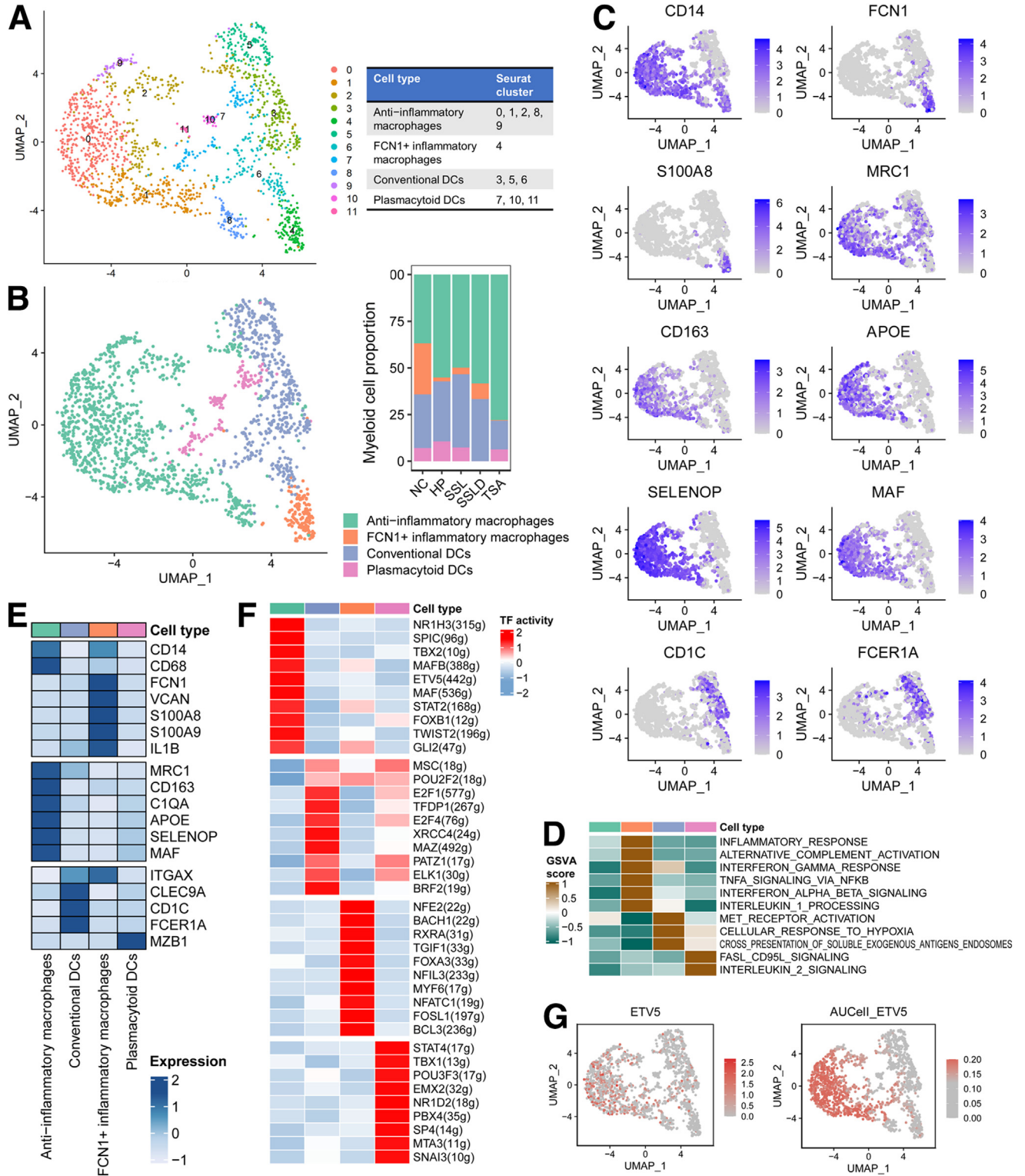
Figure 13. Myeloid cell subclusters in SLs. *A,* UMAP plot showing Seurat clusters of myeloid cells. *B,* UMAP plot for myeloid cells, color-coded for 4 subclusters (*left*) and proportion of myeloid cell types in NC and SLs (*right*). *C,* Feature plot of markers for each myeloid cell subcluster. *D,* Difference in pathway activity calculated by GSVA per cell between different myeloid cell types. *E,* Heatmap showing the canonical markers of each myeloid cell cluster. *F,* Heatmap representing the TF activity in each myeloid cell subtype. The TF activity was scored using AUCell. *G,* UMAP plot of myeloid cells color coded according to the expression of ETV5 and the area under the recovery curve (AUC) of the estimated regulon activity of ETV5, which corresponded to the degree of expression regulation of its target genes. ***P < .001.
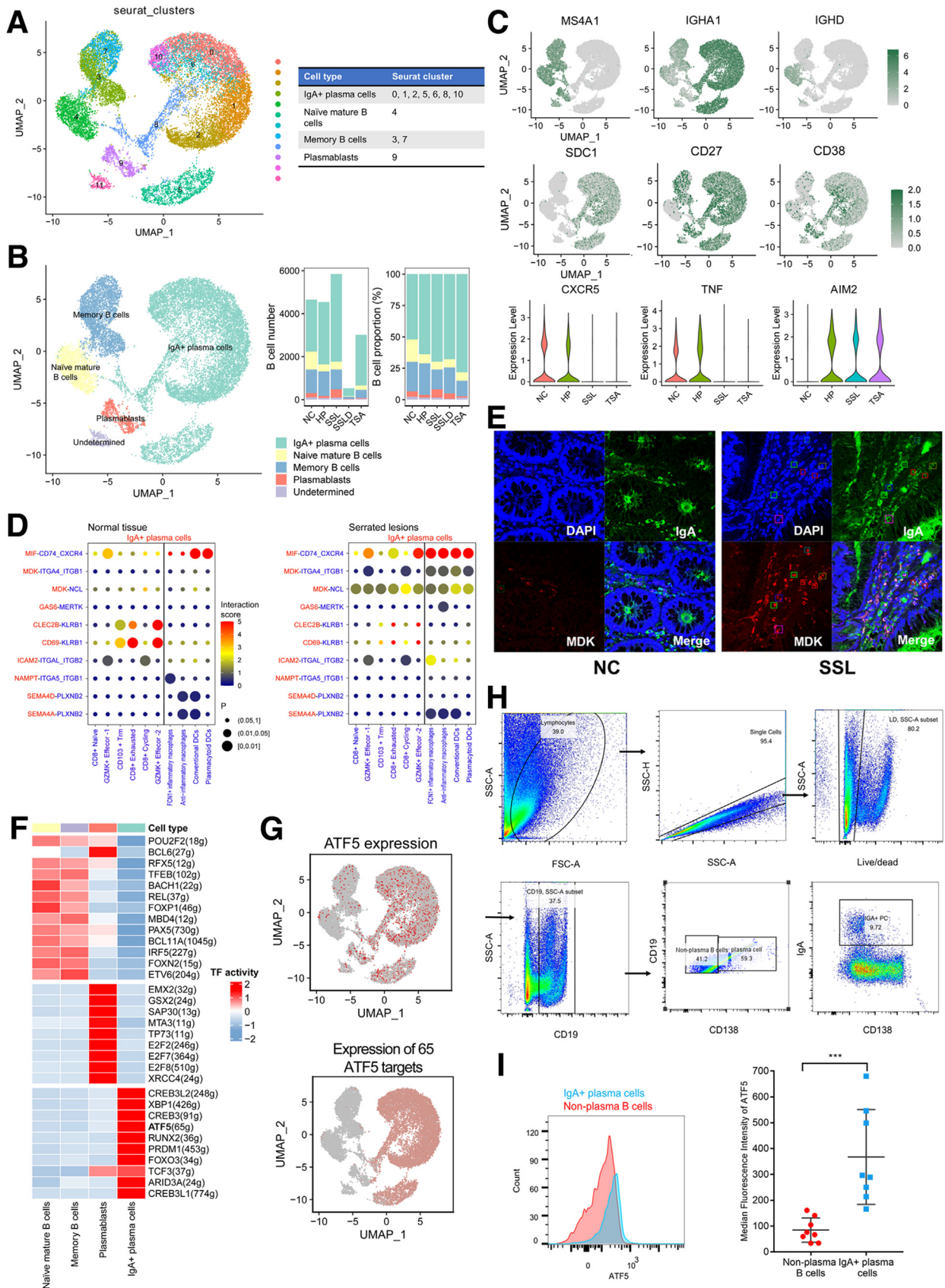
Among the fibroblast populations, we observed progressive accumulation of PDGFRA[+] fibroblasts with the disease spectrum of the serrated pathway, from 2.67% in normal colon, 13.46% in HP, 23.32% of SSL, 28.84% in TSA, to 32.05% in SSL with dysplasia. PDGFRA[+] fibroblasts featured high expression of PDGFRA, POSTN, BMP5, WNT5A, and WNT5B, indicating activation of WNT and BMP pathways, which play roles in regulating the proliferation and function of colon stem cells.[35] Immunohistochemistry analysis confirmed an increased number of PDGFRA[+] stromal cells along the 'NC-HP-SSL-SSLD' sequence (Figure 15, E).

Interestingly, consistent with these findings, a previous study in HBUS mouse model of serrated polyps demonstrated that PDGFRA[+] fibroblasts contributed to serrated polyp development by increasing MMP3 expression and facilitating membrane-bound HBEGF cleavage.[36] However, in our human samples, we found these fibroblasts did not secrete MMP3 (Figure 15, F and Figure 16, A). We compared the gene expression profile between PDGFRA[+] fibroblasts and the other fibroblasts (the full list of DEGs is summarized in Supplementary Table 15) and found that MMP11 was highly expressed in PDGFRA[+] fibroblasts (Figure 15, F and Figure 16, A). In other words, different from observations in mouse model, PDGFRA[+] fibroblasts secrete MMP11 to promote HBEGF cleavage and SL development in human serrated polyps. PDGFRA[+] fibroblasts also displayed high expression of periostin (POSTN), a secreted extracellular matrix protein that was shown to correlate with immunosuppressive premetastatic niche formation and tumor progression.[37] Considering the SLs can eventually progress to serrated CRC featuring somatic driver mutations in BRAF,[38] we investigated whether these portions of fibroblasts also existed in serrated carcinoma using publicly available data[39] of three right-sided BRAF-mutant CRCs. As expected, similar PDGFRA[+] subpopulation was also found in BRAF-mutant CRCs, which co-expressed high level of MMP11 and POSTN (see Figure 16, B).

Moreover, Stromal-2 and Stromal-3 cell proportions were also increased in the 3 types of serrated adenomas, whereas Stromal-1 cell proportion decreased across the SL spectrum. We estimated the interaction scores of prominent CXCL12-CXCR4 axis between various stromal cells and immune cells and found the strongest interactions with Stromal-1 cluster, which predominantly existed in normal tissues. The interaction was weakened in PDGFRA[+] fibroblasts, suggesting attenuated chemotaxis function of fibroblasts in SLs (Figure 16, C).

Because metabolism of EC has recently been recognized as a driver of tumor progression and angiogenesis,[40] we investigated the metabolic pathway alterations of ECs in SLs. Compared with ECs in normal samples, ECs in SL samples showed significant enrichment of oxidative phosphorylation/TCA cycle and fatty acid metabolism pathways (Figure 16, D). Active citric acid cycle and fatty acid metabolism may provide more energy for biomass synthesis and angiogenesis in proliferating tumor ECs.

Together, early molecular and immune alterations underlying the serrated pathway inferred from our scRNA-seq are illustrated in Figure 16, E.

Discussion

In the scRNA-seq landscape of serrated pathway to CRC, we concentrated on several driving questions, including: (1) what is the major characteristic of SL epithelia compared with normal colonocytes; (2) what is the heterogeneity of epithelia in SSL and TSA; (3) what is about the immune contexture in SL, and are there any metabolic factors that regulate the disease-specific tissue T cells; and (4) do any cell subpopulations expand in the microenvironment as the disease progresses to SSLD?

We found obvious antioxidant responses in the transformed epithelium of premalignant colon serrated neoplasms. Notably, the cellular responses to oxidative stress were markedly higher in SLs and BRAF-mutant CRCs, compared with conventional adenomas and BRAF wild-type CRCs. These findings suggested that redox imbalance might play a distinct role in serrated tumorigenesis. In fact, ROS is a double-edged sword in cancer initiation and progression.[41] Confronted with oncogene (BRAF)-induced oxidative stress,[42] activation of the cellular antioxidant system can be a protective mechanism that prevents cell death and promotes tumorigenesis.[43] Conversely, ROS production was shown to directly inactivate tumorigenic mutant BRAF[V600E].[44] Therefore, CRC originating from the serrated pathway could be treated differently than conventional CRC via modulation of oxidative stress. Additionally, oxidative stress-related genes could serve as biomarkers for tumors originating from the serrated pathway. We also noticed that SerpinB6 was specifically upregulated in epithelia of SLs, which has been shown to serve as protease inhibitor of cathepsin G in neutrophils.[45] In the in vitro experiments, SerpinB6 was shown to promote proliferation and migration and suppress ROS production; its exact function in colon serrated tumorigenesis remains to be elucidated.

Figure 14. (*See previous page*). **Dissection and clustering of B cells in SLs.** *A*, UMAP plot showing Seurat clusters of B cells. *B*, UMAP plot for B cells (*left*) and proportion of B cell subclusters in NC and SLs (*right*). According to classic B cell markers, we further identified naïve mature B cells (IGHD[+]), memory B cells (CD27[+]), and plasmablasts (CD38[+]). *C*, Feature plot of markers for each B cell subcluster (*upper*) and violin plot showing different expression level of chemokine and cytokine in naïve mature and memory B cells between NC and SL samples (*lower*). *D*, Network of potential interactions between IgA[+] plasma cells and myeloid or CD8[+] T cells. *E*, Confocal staining of IgA and MDK in B cells in NC and SSL tissue. *F*, Heatmap representing the TF activity in each B cell subtype. The TF activity is scored using AUCell. *G*, UMAP plots of B cells color coded according to the expression of ATF5 or the AUC of the estimated regulon activity of ATF5, which corresponded to the degree of expression regulation of its target genes. *H*, Gating strategy for FCM validation of IgA[+] PC and non-plasma B cells in SL. *I*, Validation of higher ATF5 expression in IgA[+] plasma cells in SSL tissues by FCM.
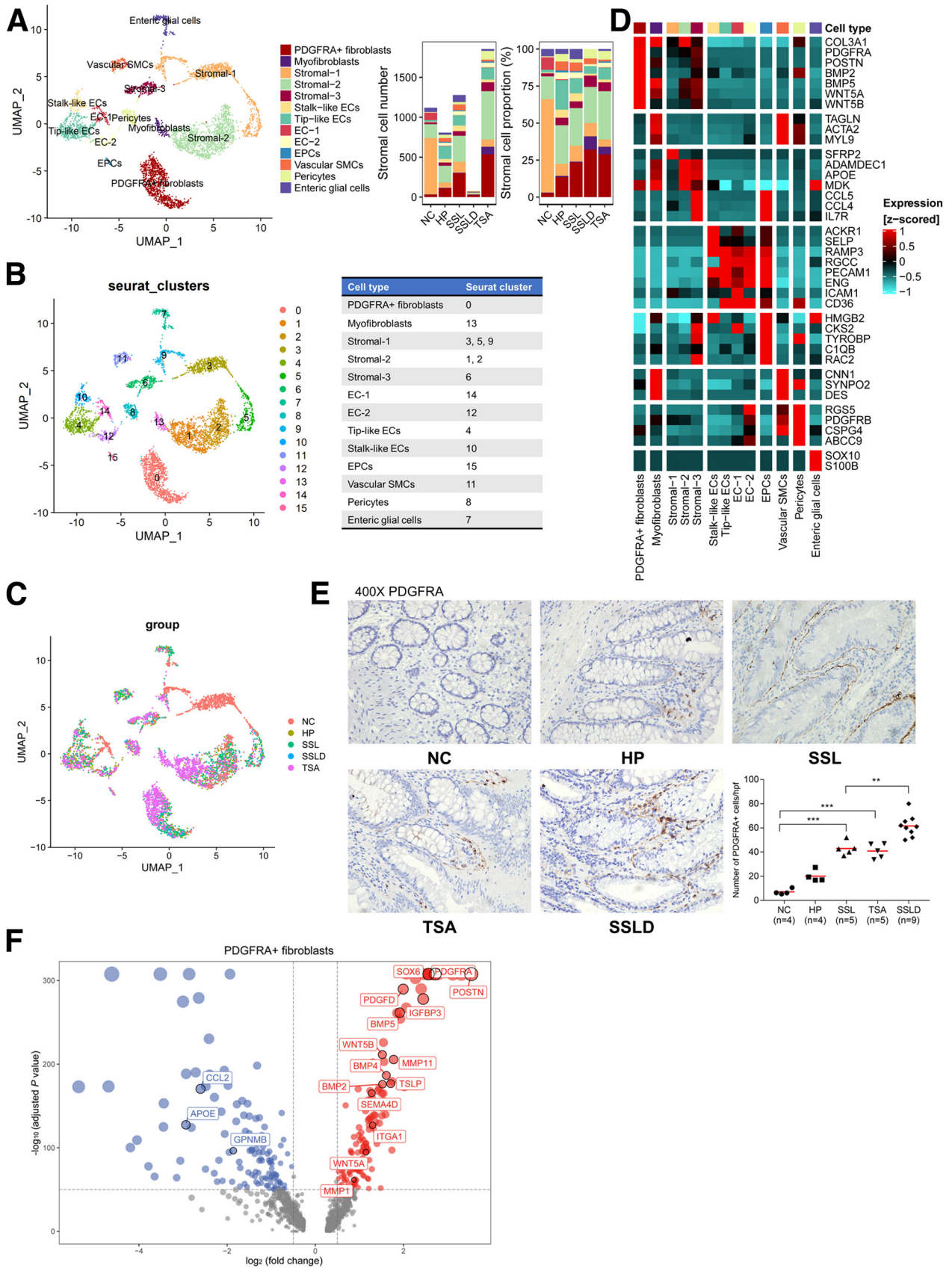
SSL and TSA originate from 2 types of epithelium showing both similar and different molecular alterations, suggesting therapeutic interventions could be developed based on molecular subtypes of SLs. In epithelium of SSL, we observed aberrant overexpression of a direct intracellular marker of proliferation (Ki67) along with Notch signaling activation. Notch signaling, which maintains stemness by blocking differentiation of intestinal stem cells to secretory lineages,[46] was found associated with SSL formation and SSLD progression in our study. Consistent with previous experimental studies in both Braf$^{V637E/+}$ and Kras^{G12D} mouse models, Notch signaling activation has been shown to play a vital role in the progression and invasion in serrated colon tumorigenesis.[47,48] For serrated CRC developed from SSL, γ-secretase inhibitor may have potential therapeutic value. Different from SSL epithelium which showed gastric metaplasia,[49] we discovered the TSA epithelium manifested a trend of Paneth cell differentiation. The abundant lysosome secretion in TSA may have complex interaction with mucosal microbiota,[50] and participate in TSA-associated carcinogenesis. Interestingly, a small fraction of Epi-SL cells was observed in normal tissue; These cells featuring high expression of MIK67 and OLFM4, suggesting their stem-cell like properties and proliferation potential. They might represent precursors of transformed cells as well as stem cells. In fact, a small amount of transformed cells can also exist in healthy tissue, but this population can be restrained and removed by immune system surveillance as well as epithelial defense under healthy conditions.[51,52]

Our data showed increased CD8$^+$ T cell activation in the entire spectrum of SLs (HP, SSL, SSLD, and TSA), which was confirmed by a recent scRNA-seq research highlighting cytotoxic immunity in SSL (MSS status) prior to MLH1 silencing and neoantigen exposure that usually occurred after the SSLD stage.[53] Furthermore, our study is the first to point out that the cytotoxic immunity in SLs resulted from increased fraction of CD103$^+$ Trm cells expressing GZMB and IFNG. In CRC, CD103$^+$ Trm was demonstrated to be associated with better prognosis and represented a large portion of CD8$^+$ lymphocytes in cancerous tissue but not in normal mucosa.[54] Consistent with our findings that CD103$^+$ Trm was more enriched in colonic neoplasms originated from serrated pathway, increased number of CD103$^+$ Trm cells was found in BRAF-mutant tumors.[55] We also observed that RA metabolism might be involved in CD103$^+$ Trm augmentation in serrated tumors, which could be a target of chemoprevention. Moreover, CD8$^+$ CD103$^+$ T cells were reported to predict response to PD-L1 blockade therapy in several studies,[24,56] suggesting that intratumor CD103$^+$ Trm component may have clinical implications for patient-selection for immunotherapy independent of MSI status.[57] Patients with serrated CRC, including at least a portion of BRAF-mutant CRC but with MSS, might be more likely to benefit from immunotherapy due to ample infiltration of Trm. This hypothesis was justified by a recent study in which the response rate of combined BRAF inhibitor and immunotherapy reached 42% in 12 patients with BRAF-mutant CRC with MSS.[58] The mechanistic details of immunometabolism in SLs and the clinical application of immunotherapy in Trm-abundant serrated CRC warrant further well-designed studies.

In previous studies, anti-tumor T cell responses against CRC were shown to increase at the early stage but decrease with tumor progression.[59] In the present study, anti-inflammatory macrophages and regulatory T cells that attenuate the anti-tumor immune response[60] were already increased, even in the premalignant stage. The imbalanced features of the TME may have implications for development of therapies targeting serrated colon tumors by preserving the cytotoxic function of Trm while suppressing the immunosuppressive cells.

Stromal components can also play vital roles in serrated tumor formation.[36] In this study, we demonstrated PDGFRA$^+$ fibroblasts enrichment in the microenvironment of all kinds of human serrated tumors, most evident in SSLD. Thus, PDGFRA expression in the stromal may serve as a biomarker for progression from SSL to SSLD. These findings were in line with previous experimental evidence of pathological PDGFRA$^+$ fibroblasts in a mouse model of HP. These pathological PDGFRA$^+$ stromal cells specifically expressed MMP11 and POSTN, which might activate EGFR signaling by promoting cleavage of HBEGF and promote tumor metastasis by facilitating immunosuppressive microenvironment, which represents a candidate intervention target for serrated CRC.

Some limitations of our study should be considered while interpreting the results. First, SSLD samples are uncommon in clinical practice, which is partly attributable to the rapid transformation of SSL to invasive cancer; therefore, we could only obtain 2 samples in this group, although we validated the findings in more slides of SSLD by immunohistochemistry. Second, considering limited tissue material of SL available (<1 cm in size), mutation status was not tested in our study, although a number of studies have shown 60% to 80% of SSL harbor BRAF mutation and approximately 50% of TSA harbor KRAS mutation.[61] Finally, in-depth mechanistic studies need to be carried out in preclinical models of SL to validate and transform our discoveries.

Figure 15. (*See previous page*). **Accumulation of PDGFRA$^+$ fibroblasts in SLs.** *A*, UMAP plot of 5,139 stromal cells in serrated lesions, color-coded by sub-cell type (*left*). Comparison of the average cell number and proportion of sub-cell types in normal mucosa and SLs (*right*) was shown. *B*, UMAP plot showing Seurat clusters of stromal cells in SLs. *C*, UMAP plot showing group distribution of stromal cells in SLs. *D*, Heatmap of canonical marker genes for fibroblasts and ECs. The color of each square indicates the z-scaled average gene expression (*cyan to red*). *E*, Immunohistochemistry staining and quantitative analysis of PDGFRA$^+$ cells in the stromal of normal, HP, SSL, TSA, and SSLD tissues. *F*, Transcriptome differences between PDGFRA$^+$ fibroblasts and other fibroblasts. **$P < .01$; ***$P < .001$.

In summary, our study unveiled the early molecular traits and immune responses across the 4 classes of SLs at single-cell resolution. Our findings help enhance the understanding of the pathogenesis of serrated neoplasia pathway and provide insights for further investigation of novel therapeutic targets.
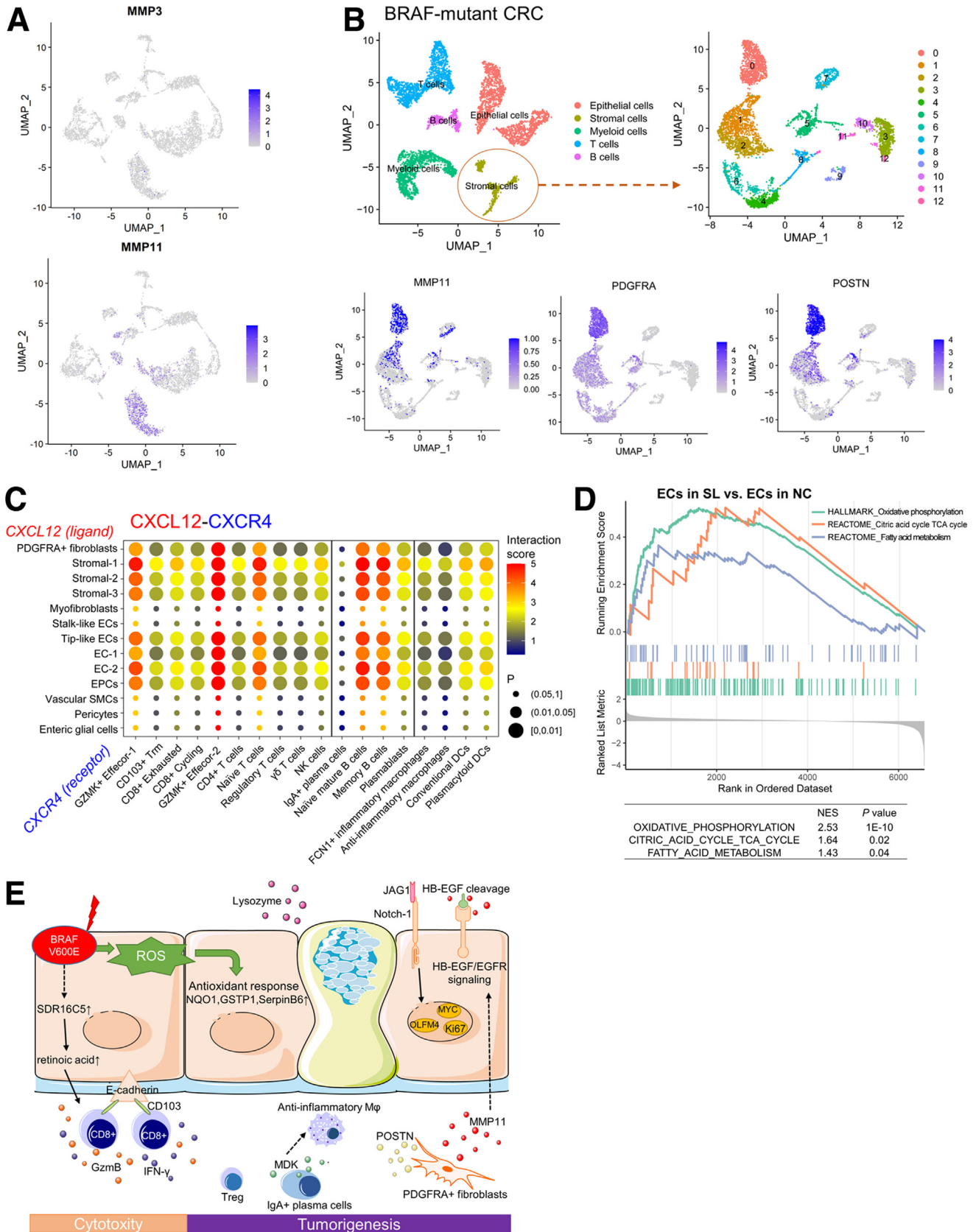
Materials and Methods

Patient Samples

We obtained 19 human colorectal samples from Renji Hospital, School of Medicine, Shanghai Jiao Tong University for scRNA-seq. Patient clinicopathological data and endoscopy reports of lesion size were recorded. Each participant has provided written informed consent. Ethical approval was obtained from Shanghai Renji Hospital Ethics Committee (KY2021-004).

Single-cell Suspensions for SL Samples

Fresh serrated polyp tissues resected from colonoscopy were promptly washed with phosphate buffered saline (PBS) and divided into 2 parts. One part of specimen was fixed in formalin and sent for pathological evaluation. Formalin-fixed paraffin-embedded (FFPE) biopsy tissue with hematoxylin-eosin staining were interpreted by an experienced gastrointestinal pathologist who was blinded to patients' clinical and endoscopic data. The pathological diagnosis of HP, SSL, SSLD, and TSA were according to the 2019 World Health Organization classification for serrated lesions (fifth edition).[62] Another part of tissue was stored in MACS Tissue Storage Solution (cat#130-100-008, Miltenyi Biotec) at 4 °C, which was then used to generation of single-cell suspension within 24 hours.

To disassociate the serrated polyp tissue, we collected the tissue in a 6-cm plate added with RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS) and kept on ice. Then the tissue was cut into small pieces of about 1 mm^2 and transferred into a 1.5-mL EP tube. We added complete media (supplemented with 10% FBS) and resuspended to 1 mL, followed by adding 20 μg/mL DNaseI (D8071-100mg, Solarbo), and 200 μg/mL TL enzyme (Liberase TL Research Grade, cat#5401020001, Sigma). The tube was placed on a preheated rocking bed, shaking at 250 rpm at 37 °C, and digested for 1 hour. We turned the tube upside down several times every 15 minutes. The resulting suspension was added with an additional 200 uL FBS and passed through a 70-μm sieve. Then, it was subjected to centrifugation at 500g, 4 °C for 5 minutes. The supernatant was discarded, and then the precipitate was lysed the with 1 mL Red Blood Cell Lysis Buffer (cat#00-4300-54, Invitrogen) for 10 minutes at room temperature. After washing cells with PBS twice (centrifugation at 600g, 4 °C for 5 minutes), we collected cells form the bottom and resuspended to 500 uL with PBS. Then, single-cell suspension was filtered with a 40-μm sieve and counted using a cell count plate.

Library Preparation and Sequencing

Cells were counted, and cell density was adjusted to that recommended for 10× Genomics Chromium single-cell 3' v3 processing and library preparation. Sequencing was performed on an Illumina platform (NovaSeq 6000), by GENERGY BIO (Shanghai, China), at a sequencing depth of approximately 50,000 reads per single cell. The sequencing data were deposited at the Genome Sequence Archive database under the accession number HRA002611.

Single-cell Gene Expression Quantification

Single-cell 3' libraries were constructed using a commercial 10× Genomics platform (10× Genomics, Pleasanton, CA). Single-cell transcriptomic amplification and complementary deoxyribonucleic acid (cDNA) library preparation were performed by GENERGY BIO (Shanghai, China) using 10× Genomics Chromium single-cell 3' v3 according to the manufacturer's instructions. The cDNA libraries were sequenced on NovaSeq 6000 System (Illumina, San Diego, CA), at a sequencing depth of about 50,000 reads per single cell.

Pre-processing of scRNA-seq sequencing data was conducted with Cell Ranger (version 3.0.2, 10× Genomics) using default parameters. The raw sequencing data were demultiplexed from raw base call (BCL) files and converted into fastq files using the 'mkfastq' command. The fastq files were aligned to the Genome Reference Consortium Human Build 38 (GRCh38). Counting all confidently mapped reads, filtered gene-barcode matrices were generated using the 'count' command for each sample, with low unique molecular identifiers (UMIs) droplets removed for minimizing the number of empty droplets. Low-quality cells were removed if cells had either fewer than 1001 UMIs, over 4000 or less than 1001 expressed genes, less than 5% UMIs from ribosomal genes or over 25% UMIs derived from the mitochondrial genome. We filtered genes that were detected in less than 3 cells. Canonical correlation analysis was used to remove the batch effect from 18 patients in which every patient was assigned to one batch. We annotated cell components using a combination of reference based (R package 'SingleR', version 0.99.10) and manual annotation.

Clustering of Single-cell Data Matrix

We performed the clustering of various cell types using R package 'Seurat' (version 4.0.3). Data normalization was completed via 'NormalizeData' function, with the default scaling parameter of 10,000 and log normalization method. 'FindVariableGenes' function was employed to find 4000 genes of the highest variance. We standardized the data with the 'ScaleData' function. After performing principal component analysis using highly variable genes, the top 20 principal components and a resolution of 0.5 were selected for the following cluster analysis and visual dimensionality

Figure 16. (*See previous page*). **Stromal cell dynamics in normal colonic tissue and colorectal tumors derived from serrated pathway.** *A*, Feature plot showing expression of MMP3 and MMP11 in stromal cells of SLs. *B*, Identification of MMP11-secreting PDGFRA$^+$ fibroblasts in the stromal cells of BRAF-mutant CRCs (*upper*). Feature plot showing expression of PDGFRA, MMP11, and POSTN in stromal cells of BRAF mutated CRCs (*bottom*). Data were retrieved from Samsung Medical Center cohort reported in a previous study.[39] *C*, Dot plot showing cell-cell interaction of CXCL12-CXCR4 axis between stromal cells and immune cells. *D*, Pathway enrichment analysis of ECs in SLs vs NCs. *E*, Diagram illustrating early molecular and immune alterations underlying the serrated pathway inferred from scRNA-seq.

reduction by UMAP for dimension reduction. We used the 'FindAllMarkers' or 'FindMarkers' function to determine the marker genes of each cluster compared to all other clusters or to specific cluster(s). 'FeaturePlot' or 'VlnPlot' function was used for presentation of gene expression levels. We labeled the obtained clusters as epithelial cells, T cells, B cells, myeloid cells, or stromal cells through known classic markers (epithelial cells: EPCAM; T cells: CD3D; B cells: MS4A1, CD79A; myeloid cells: CD14, CD163; stromal cells: COL1A1, PECAM1) and finally analyzed each of these clusters separately to identify the finer clusters by repeating the above operations. Cell-cycle phases were predicted using a function included in Seurat that scores each cell based on expression of canonical marker genes for S and G2/M phases.

Copy Number Variation Estimation

The 'InferCNV' R package (version 1.8.1) was used to estimate the copy number variations (CNVs) in epithelial cells based on transcriptome profiles (inferCNV of the Trinity CTAT Project, https://github.com/broadinstitute/inferCNV).[63] The CNV scores of epithelial cells from the normal samples was as the reference cells in the final estimation of CNVs. For the inferCNV analysis, the default parameters were used: 'True' for denoise choice, 'True' for hidden Markov model setting, and '0.1' for signal cutoff. To reduce the possibility of false-positive CNV calls, the default Bayesian latent mixture model was implemented to identify the posterior probabilities of alterations in each cell. Low-probability CNVs were filtered using the default value of '0.5' for the threshold. To determine the clonal CNV changes in each SL sample, the 'subcluster' method was utilized on the CNVs generated by the hidden Markov model. GRCh38 cytoband information was used to convert each CNV to a p- or q- arm level change for simplification based on its location. Each CNV was annotated to be either a gain or a loss. The CNV score of each cell was calculated as quadratic sum of CNV region.

Differentially Expressed Gene and Pathway Analysis

To gain a biological understanding of the identified DEGs, we conducted GO and KEGG analyses, which were conducted through the 'clusterProfiler' R package (version 3.18.1) with the gene set background from Molecular Signatures Database (version 7.4).[64] Volcano plot for DEGs were generated by 'ggplot2' R package (version 3.3.3). In single-cell sequencing analysis, DEGs between cell clusters were defined as those with adjusted $P < .05$ and $|\log_2 FC| > 0.25$.

Gene Set Variation Analysis Among Different Cell Types

A subset-enriched term of biological relevance was represented as enrichment score using the corresponding genes or markers and was calculated by the gene set variation analysis approach through using 'GSVA' R package

(version 1.36.2).[65] Heatmap for biological pathways of interest were generated by 'ComplexHeatmap' R package (version 2.5.5). Evaluation of metabolic activity per single cell was performed using the recently-proposed 'scMetabolism' R package, based on VISION, AUCell, and ssGSEA methods.[28]

Trajectory Analysis

We explored the epithelial cell transition from transit amplifying cells and intermediate cells to SL-specific cells (Epi-1 and Epi-2) using R package 'Monocle' (version 2.20.0). The standard workflow was employed to identify high variance gene between these epithelial cell subtypes using 'dispersionTable' function. Next, the function 'reduceDimension' with the 'DDRTree' method was employed to reduce the dimensions, and the cells were then ordered based on the predicted pseudotime via 'orderCells' function.

SCENIC Analysis

pySCENIC (version 0.11.0) was used to identify gene regulatory networks and key TFs in scRNA-seq datasets using default parameters. In the first step, given a pre-defined list of TFs (hs_hgnc_tfs.txt, https://github.com/aertslab/pySCENIC/tree/master/resources), co-expression modules between these TFs and putative target genes were inferred using Arboreto algorithm (GRNBoost) from normalized expression, followed by pruning the insignificant co-expression pair. Secondly, the function 'ctx' will measure the enrichment of the TFs motif in putative regulatory regions of the target gene based on the cisTarget databases and the motif annotation tables (https://resources.aertslab.org/cistarget/). The TF and its predicted targets will be retained if the motif of the TF is significantly enriched in one of its modules. Finally, the activity score of the predicted regulons in the individual cells was calculated using AUCell algorithm, which measures the enrichment of target genes in whole-genome ranking.

To measure the specificity of TFs for clusters, the regulon specificity score was calculated as the Jensen-Shannon divergence between the distribution of AUCell score and clusters. TFs with top regulon specificity score will be kept as cluster-specific TFs.

We predicted and visualized regulon networks in epithelial cell subclusters, as reported in a recent study.[53] The cisTarget step in SCENIC produced a list of the TF and all its target genes in each regulon, with the prediction importance from GRNBoost as the corresponding weight of each TF-target pair. The list was transformed to a target-by-TF weight matrix via 'recast' function, which characterized each regulon based on its transcriptional weights on all its targets. Based on the Euclidean distance of weight matrix, k nearest neighbors of each regulon were found, with k selected as the square root of the number of regulons. TFs with similar transcription regulation profile (target genes and corresponding weight) was grouped together using laiden algorithm at the resolution_parameter of 2. In each laiden cluster, the adjacency between each pair of regulon was set as the weight of the TF-target pair if there is a TF-

target relationship between them, generating a transcription regulatory network. The network is visualized using R package 'igraph' (version 1.2.10), with the color of vertices as mean scaled AUCell score and the width of edges as adjacency between regulons in this cluster.

Receptor-ligand Pairing Analysis

To qualify the cell-cell communication differences, we applied R package 'CellChat' (version 1.1.0) to infer and analyze the intercellular communication network on normal and SL samples separately. Briefly, a CellChat object was created from raw count matrix and cell-type annotation. Inferred interactions were filtered using 'subsetCommunication' function with the default parameter. The strength of cell interactions was quantitated as the mean log-normalized count of the ligand gene in the source cell type and receptor gene(s) in the target cell type.

Public Database Analysis

We obtained external bulk RNA-seq data from GSE76987 (right-sided colon tissues of 10 normal, 10 conventional polyps, 21 SSLs), EMTAB7960 (22 SSL tissues and 20 SSL with dysplasia), and GSE39582 (51 patients with BRAF-mutant CRC, 217 patients with KRAS-mutant CRC, and 257 patients with wild-type CRC). TCGA (http://cancergenome.nih.gov) CRC data were obtained by R package 'TCGAbiolinks' (version 2.20.1).

For integrated analysis of TCGA and GEO, the transcriptome profiles of patients with COAD and READ from the TCGA cohort were downloaded using R package TCGA-biolink and transformed as logarithm 'Transcripts Per Kilobase Million (TPM)'. The expression profiling by microarray from GSE39582 was log-transformed. The samples were removed if they did not have matched information about BRAF mutation or microsatellite instability. And then an integrated analysis of the 2 cohorts was conducted with batch effect removed by the function 'combat' in R package 'sva' (version 3.40.0). We accessed the scRNA-seq data of 3 right-sided BRAF-mutant CRCs (SMC-03, 10, 17) from Samsung Medical Center dataset in a previous study.[39]

Immunohistochemistry and Confocal Staining

FFPE tissue blocks of 4 normal colon tissue, 4 HPs, 8 SSLs, 5 TSAs, 4 TAs, 9 SSLs with cytological dysplasia, and 7 BRAF-mutant right-sided CRCs from Shanghai Renji Hospital were used for immunohistochemistry staining. All FFPE colonic samples were stained with hematoxylin and eosin for histopathological evaluation at the Shanghai Renji Hospital. For immunohistochemistry staining, HP, SSL, and SSLD samples were collected from proximal colon, NCs were collected from ascending (n = 1), transverse (n = 1), and descending colon (n = 2), whereas TSA samples were collected from distant colon or rectum. In brief, after antigen retrieval, sections of colonic samples were blocked with goat serum for 30 minutes. Then, the sections were incubated overnight with a primary antibody against GSTP1 (1:250, ab138491, Abcam), NQO1 (1:1000, ab28947, Abcam), DUOX2 (1:1000, NB110-6157655, NOVUS), 8-OHdG

(1:100, ab48508, Abcam), SERPINB6 (1:200, 14962-1-AP, Proteintech), JAG1 (1:200, ab109536, Abcam), MYC (1:100, ab32072, Abcam), FOXP3 (1:200, ab4728, Abcam), ITGAE (1:500, ab224202, Abcam), and PDGFRA (1:500, ab134123, Abcam) at 4 °C followed by incubation with a horseradish-peroxidase-conjugated secondary antibody at room temperature for 30 minutes. Then, 3,3-diaminodbenzidine substrate was added. Five random fields were analyzed under a light microscope.

Immunofluorescence staining was performed on human SL tissues. Briefly, NC, SSL, and TSA tissues were fixed in 4% paraformaldehyde, paraffin-embedded, and sectioned. Slides were washed 3 times in PBS, followed by blocking with 10% goat serum for 20 minutes. Sections were then boiled in sodium citrate solution to retrieve antigens. Then, sections were incubated overnight with primary antibodies against MUC2 (1:200, ab231427, Abcam) and LYZ (1:250, ab108508, Abcam); CD8 (1:50, ab33786, Abcam) and ITGAE (1:500, ab224202, Abcam); IgA (1:100, ab124716, Abcam) and MDK (1:50, sc-46701, Santa Cruz) at 4 °C. After washing 3 times in PBS, slides were incubated with fluorescent dye-labeled secondary antibodies (1:500; Invitrogen, Carlsbad, CA) for 30 minutes at room temperature. Slides were then stained with DAPI (Southern Biotech, Birmingham, AL) after washing 3 times with PBS. Images were captured under a confocal fluorescence microscope (LSM-710; Carl Zeiss, Jena, Germany).

Cell Line Experiments

RKO cells were obtained from American Type Culture Collection and were cultured in RPMI1640 medium (Gibco) supplemented with 10% FBS. The cell line was verified by short tandem repeat profile analysis. For siRNA transfection, RKO cells were transfected with 50 pmol siRNA (siNC or siSERPINB6; TSINGKE Biotechnology, Shanghai, China) for 6 hours, then used for different assays after 48 hours. The sense and antisense sequences of the SERPINB6 siRNA are as follows: 5'-CAAAUCUUGGUGCUUCCAUAU -3' and 5'-AUAUGGAAGCACCAAGAUUUG -3', respectively. For SerpinB6 rescue experiment, RKO cells were transfected with FuGENE HD transfection reagent (Promega, Madison, WI) after 4 hours of siSERPINB6 transfection. We transfected 0.6 ug SerpinB6 plasmid DNA with 0.6 uL FuGENE Reagent per 6-well plate for another 2 hours. Transfection medium was replaced by cell growth medium 6 hours after transfection. Proteins from cells were extracted as previously described.[66] The follow primary antibodies were used for Western blotting: anti-SERPINB6 (14962-1-AP, Proteintech, IL) and anti-GAPDH (KC-5G5, Kangchen Company, Shanghai, China).

Cell proliferation was measured by EdU assay kit (C10338-1, RiboBio, Guangdong, China) according to the manufacturer's instructions. For transwell assay, a total of 2×10^5 transfected RKO cells were seeded into the upper chamber of a 24-well polycarbonate transwell filter (8 μm pore size, Corning Incorporated, Corning, NY). RPMI1640 media containing 30% FBS was placed in the lower chamber. After 48 hours of incubation, the migration cells adhering to the lower surface were fixed with 4%

paraformaldehyde and stained with 0.1% crystal violet. The area of positive staining was calculated in 5 random fields using Image-Pro Plus software (version 6.0, Media Cybernetics). ROS levels were detected using Reactive Oxygen Species Assay Kit (S0033, Beyotime, China) with DCFH-DA probe. Briefly, DCFH-DA was diluted at 1:1000 with serum-free RPMI1640 medium and incubated at 37 °C for 20 minutes. Then, cells were washed in serum-free RPMI1640 medium and collected for FCM analysis. N-acetyl L-cysteine (ST1546, Beyotime, China) was used as a ROS scavenger.

Quantitative Real-time Polymerase Chain Reaction

Total RNA was extracted from 8 SSL samples with paired normal tissues collected from Shanghai Renji Hospital, using TRIzol reagent (Invitrogen, Carlsbad, CA). The cDNA was synthesized using PrimeScript RT reagent Kit (Takara, Japan). Quantitative RT-PCR was performed using SYBR-Green Master mix (Takara, Japan) according to the manufacturer's instructions. The expression levels of genes were analyzed with GAPDH serving as an internal control. The primer sequences used were as follows: SERPINB6 (forward primer: AGGGAAACACCGCTGCACAGAT; reserve primer: GTGCCAGTCTTGTTCACTTCGG); SDR16C5 (forward primer: TGCACGCCTATACCTGCGATTG; reserve primer: GGCATTGTTGATTAGGATGGAAAC); GAPDH (forward primer: GTCTCCTCTGACTTCAACAGCG; reserve primer: ACCACCCTGTTGCTGTAGCCAA).

Immune Cell Isolation of Human Colon Adenomas

To validate the findings of T and B cells in the SL microenvironment, we obtained single cell suspension of another 8 fresh SSL tissues resected from colonoscopy resection, as described in the first part of the Supplementary Methods. The histological diagnosis was confirmed by 2 experienced pathologists blinded to the endoscopic imaging independently. To isolate immune cells, after filtered through a 70-μm sieve and centrifugation, the cells at the bottom were resuspended and then overlaid onto 33% Percoll (GE Healthcare, Chicago, IL). The suspension was then centrifuged at 900 g for 30 minutes at room temperature with the brake off. After the spin, red blood cells were lysed as previously described. The isolated leukocytes were then washed twice before staining.

In Vitro Generation of Human CD8$^+$CD69$^+$CD103$^+$ T Cells

We used an in-vitro model to generate CD103$^+$ Trm cells from human human peripheral blood mononuclear cells (PBMCs, as previously reported.[67] In brief, PBMCs were cultured in 24-well plates at 5×10^5 cells /mL in RPMI1640 medium supplemented with 10% heat-inactivated FBS, 1% penicillin/streptomycin (Gibco), and 50 mM 2-mercaptoethanol (Invitrogen). To enrich CD103$^+$ Trm cells, we added 2 ng/mL rhIL-2 (PeproTech, Cranbury, NJ) for the whole course, and added rhIL-15 (50 ng/mL; R&D

Systems, Minneapolis, MN) for 3 days, followed by a further 3 days of rhTGF-β (50 ng/mL; PeproTech). The medium was refreshed on day 3. RA (HY-14649, MedChemExpress) at the concentration of 10 nM and 100 nM was used for investigating the effects on the expansion of CD8$^+$CD69$^+$CD103$^+$ T cells. The cells were collected for FCM analysis on day 6 (see Figure 12A).

Flow Cytometric Analysis

Cells were harvested and washed with FACS buffer. For surface staining, the 1×10^6 cells in 100-μL single-cell suspension were stained with antibodies for 30 minutes in the dark at 4 °C. The following antibodies were used: live and dead cell stain FVS (BD Biosciences), anti-CD3, anti-CD4, anti-CD8, anti-CD103/ITGAE, anti-CD19, anti-CD27, anti-IgD, anti-CD38, and anti-CD138 (BD Biosciences, San Jose, CA); and anti-CD69, anti-CD18/ITGB2, anti-KLRG1 (BioLegend, San Diego, CA). For intracellular cytokine staining, immune cells were incubated in complete RPIM1640 containing 10% FBS and leukocytes activation cocktail with GolgiPlug (BD Biosciences) at 37 °C for 5 hours, and then surface-stained cells were fixed and permeabilized with Cytofix/Cytoperm solution (BD Biosciences) for 20 minutes at 4 °C. Subsequently, cells were stained with anti-GZMB, anti-GZMK, and anti-IgA (BD Biosciences) for 30 minutes at 4 °C. For staining TFs, cells were fixed and permeabilized with Transcription Factor Buffer Set (BD Biosciences). Briefly, after surface staining, cells were fixed and permeabilized for 50 minutes at 4 °C, then washed and stained with ATF5 (Abcam, Cambridge, UK) for 50 minutes at 4 °C. FCM was carried out using Celesta (BD Biosciences) and analyzed using FlowJo software (version 10.6.2, Tree Star).

Survival Analysis

We explored the prognostic relevance of CD8$^+$ T cell types, using signature genes derived from our scRNA-seq data. The patient data used in survival analysis were obtained from GSE39582, incorporating 263 patients with CRC with BRAF or KRAS mutation and available survival information. The selected criteria of signature genes for one cell type were DEGs detected more than 25% of the target cell type, with adjusted P value < .05 and the threshold of $\log_2$(fold change) at 0.8, according to previous studies.[68] For CD103$^+$ Trm, a predefined Trm CD8$^+$ gene signature by Savas et al was used.[23] The survival analysis was performed by Cox regression model, using recurrence-free survival as the outcome. The analysis was done by R package 'survival' (version 3.2-3) and 'survminer' (version 0.4.8).

Statistical Analysis

All statistical analyses were performed using R version 4.1.0 (https://www.r-project.org/), Python version 3.7.10, and Graphpad version 7.0. For all statistical tests, a 2-sided P value less than .05 was regarded as statistically significant. Comparisons between 2 groups were performed by the unpaired Student t test. Data were expressed as mean ± standard deviations. Scripts and code for the analysis are available at Github (https://github.com/xlucpu/scRNAseq-SL).

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69:7–34.
2. Ijspeert JEG, Vermeulen L, Meijer GA, Dekker E. Serrated neoplasia-role in colorectal carcinogenesis and clinical implications. Nat Rev Gastroenterol Hepatol 2015; 12:401–409.
3. Ahadi M, Sokolova A, Brown I, Chou A, Gill AJ. The 2019 World Health Organization classification of appendiceal, colorectal and anal canal tumours: an update and critical assessment. Pathology 2021;53:454–461.
4. Crockett SD, Snover DC, Ahnen DJ, Baron JA. Sessile serrated adenomas: an evidence-based guide to management. Clin Gastroenterol Hepatol 2015;13:11–26.e1.
5. De Palma FDE, D'Argenio V, Pol J, Kroemer G, Maiuri MC, Salvatore F. The molecular hallmarks of the serrated pathway in colorectal cancer. Cancers (Basel) 2019;11:1017.
6. Tsai J-H, Jeng Y-M, Yuan C-T, Lin Y-L, Cheng M-L, Liau J-Y. Traditional serrated pathway-associated colorectal carcinoma: morphologic reappraisal of serrated morphology, tumor budding, and identification of frequent PTEN alterations. Am J Surg Path 2019; 43:1042–1051.
7. Oono Y, Fu K, Nakamura H, Iriguchi Y, Yamamura A, Tomino Y, Oda J, Mizutani M, Takayanagi S, Kishi D, Shinohara T, Yamada K, Matumoto J, Imamura K. Progression of a sessile serrated adenoma to an early invasive cancer within 8 months. Dig Dis Sci 2009;54:906–909.
8. Amemori S, Yamano HO, Tanaka Y, Yoshikawa K, Matsushita HO, Takagi R, Harada E, Yoshida Y, Tsuda K, Kato B, Tamura E, Eizuka M, Sugai T, Adachi Y, Yamamoto E, Suzuki H, Nakase H. Sessile serrated adenoma/polyp showed rapid malignant transformation in the final 13 months. Dig Endosc 2020;32:979–983.
9. Monreal-Robles R, Jáquez-Quintana JO, Benavides-Salgado DE, González-González JA. Serrated polyps of the colon and rectum: a concise review. Rev Gastroenterol Mex (Engl Ed) 2021;86:276–286.
10. Gupta V, East JE. Optimal endoscopic treatment and surveillance of serrated polyps. Gut Liver 2020; 14:423–429.
11. Lee YM, Huh KC. Clinical and biological features of interval colorectal cancer. Clin Endosc 2017;50:254–260.
12. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. Gastroenterology 2010; 138:2088–2100.
13. Pai RK, Bettington M, Srivastava A, Rosty C. An update on the morphology and molecular pathology of serrated colorectal polyps and associated carcinomas. Mod Pathol 2019;32:1390–1415.
14. Parikh K, Antanaviciute A, Fawkner-Corbett D, Jagielowicz M, Aulicino A, Lagerholm C, Davis S, Kinchen J, Chen HH, Alham NK, Ashley N, Johnson E, Hublitz P, Bao L, Lukomska J, Andev RS, Björklund E, Kessler BM, Fischer R, Goldin R, Koohy H, Simmons A. Colonic epithelial cell diversity in health and inflammatory bowel disease. Nature 2019;567:49–55.
15. Yuen ST, Wong MP, Chung LP, Chan SY, Cheung N, Ho J, Leung SY. Up-regulation of lysozyme production in colonic adenomas and adenocarcinomas. Histopathology 1998;32:126–132.
16. Pu Y, Song Y, Zhang M, Long C, Li J, Wang Y, Xu Y, Pan F, Zhao N, Zhang X, Xu Y, Cui J, Wang H, Li Y, Zhao Y, Jin D, Zhang H. GOLM1 restricts colitis and colon tumorigenesis by ensuring Notch signaling equilibrium in intestinal homeostasis. Signal Transduct Target Ther 2021;6:148.
17. Niu Y, Shao Z, Wang H, Yang J, Zhang F, Luo Y, Xu L, Ding Y, Zhao L. LASP1-S100A11 axis promotes colorectal cancer aggressiveness by modulating TGFbeta/Smad signaling. Sci Rep 2016;6:26112.
18. Feng M, Feng J, Chen W, Wang W, Wu X, Zhang J, Xu F, Lai M. Lipocalin2 suppresses metastasis of colorectal cancer by attenuating NF-kappaB-dependent activation of snail and epithelial mesenchymal transition. Mol Cancer 2016;15:77.
19. Soares de Lima Y, Arnau-Collell C, Díaz-Gay M, Bonjoch L, Franch-Expósito S, Muñoz J, Moreira L, Ocaña T, Cuatrecasas M, Herrera-Pariente C, Carballal S, Moreno L, Díaz de Bustamante A, Castells A, Bujanda L, Cubiella J, Rodríguez-Alcalde D, Balaguer F, Castellví-Bel S. Germline and somatic whole-exome sequencing identifies new candidate genes involved in familial predisposition to serrated polyposis syndrome. Cancers (Basel) 2021;13:929.
20. Sekine S, Yamashita S, Tanabe T, Hashimoto T, Yoshida H, Taniguchi H, Kojima M, Shinmura K, Saito Y, Hiraoka N, Ushijima T, Ochiai A. Frequent PTPRK-RSPO3 fusions and RNF43 mutations in colorectal traditional serrated adenoma. J Pathol 2016; 239:133–138.
21. Lutter L, Roosenboom B, Brand EC, Ter Linde JJ, Oldenburg B, van Lochem EG, Horje Talabur Horjus CS, van Wijk F. Homeostatic function and inflammatory activation of ileal CD8(+) tissue-resident T cells is dependent on mucosal location. Cell Mol Gastroenterol Hepatol 2021;12:1567–1581.
22. FitzPatrick MEB, Provine NM, Garner LC, Powell K, Amini A, Irwin SL, Ferry H, Ambrose T, Friend P, Vrakas G, Reddy S, Soilleux E, Klenerman P, Allan PJ. Human intestinal tissue-resident memory T cells comprise transcriptionally and functionally distinct subsets. Cell Rep 2021;34:108661.
23. Savas P, Virassamy B, Ye C, Salim A, Mintoff CP, Caramia F, Salgado R, Byrne DJ, Teo ZL, Dushyanthen S, Byrne A, Wein L, Luen SJ, Poliness C, Nightingale SS, Skandarajah AS, Gyorki DE, Thornton CM, Beavis PA, Fox SB, Darcy PK, Speed TP, Mackay LK, Neeson PJ, Loi S. Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. Nat Med 2018;24:986–993.
24. Banchereau R, Chitre AS, Scherl A, Wu TD, Patil NS, de Almeida P, Kadel EE Iii, Madireddi S, Au-Yeung A, Takahashi C, Chen YJ, Modrusan Z, McBride J, Nersesian R, El-Gabry EA, Robida MD, Hung JC, Kowanetz M, Zou W, McCleland M, Caplazi P, Eshgi ST, Koeppen H, Hegde PS, Mellman I, Mathews WR, Powles T, Mariathasan S, Grogan J, O'Gorman WE.

Intratumoral CD103$^+$ CD8$^+$ T cells predict response to PD-L1 blockade. J Immunother Cancer 2021;9:e002231.

25. Kirby JA, Bone M, Robertson H, Hudson M, Jones DE. The number of intraepithelial T cells decreases from ascending colon to rectum. J Clin Pathol 2003;56:158.

26. Tyler CJ, Guzman M, Lundborg LR, Yeasmin S, Perez-Jeldres T, Yarur A, Behm B, Dulai PS, Patel D, Bamias G, Rivera-Nieves J. Inherent immune cell variation within colonic segments presents challenges for clinical trial design. J Crohns Colitis 2020;14:1364–1377.

27. Flerin NC, Pinioti S, Menga A, Castegna A, Mazzone M. Impact of immunometabolism on cancer metastasis: a focus on T cells and macrophages. Cold Spring Harb Perspect Med 2020;10:a037044.

28. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, Cheng Y, Huang S, Liu Y, Jiang S, Liu J, Huang X, Wang X, Qiu S, Xu J, Xi R, Bai F, Zhou J, Fan J, Zhang X, Gao Q. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. Cancer Discovery 2022;12:134–153.

29. Cui HS, Joo SY, Cho YS, Kim JB, Seo CH. CPEB1 or CPEB4 knockdown suppresses the TAK1 and Smad signalings in THP-1 macrophage-like cells and dermal fibroblasts. Arch Biochem Biophys 2020;683:108322.

30. A-González N, Castrillo A. Liver X receptors as regulators of macrophage inflammatory and metabolic pathways. Biochim Biophys Acta 2011;1812:982–994.

31. Sharonov GV, Serebrovskaya EO, Yuzhakova DV, Britanova OV, Chudakov DM. B cells, plasma cells and antibody repertoires in the tumour microenvironment. Nat Rev Immunol 2020;20:294–307.

32. Yang M, Long D, Hu L, Zhao Z, Li Q, Guo Y, He Z, Zhao M, Lu L, Li F, Long H, Wu H, Lu Q. AIM2 deficiency in B cells ameliorates systemic lupus erythematosus by regulating Blimp-1-Bcl-6 axis-mediated B-cell differentiation. Signal Transduct Target Ther 2021;6:341.

33. Zhang Y, Zuo C, Liu L, Hu Y, Yang B, Qiu S, Li Y, Cao D, Ju Z, Ge J, Wang Q, Wang T, Bai L, Yang Y, Li G, Shao Z, Gao Y, Li Y, Bian R, Miao H, Li L, Li X, Jiang C, Yan S, Wang Z, Wang Z, Cui X, Huang W, Xiang D, Wang C, Li Q, Wu X, Gong W, Liu Y, Shao R, Liu F, Li M, Chen L, Liu Y. Single-cell RNA-sequencing atlas reveals an MDK-dependent immunosuppressive environment in ErbB pathway-mutated gallbladder cancer. J Hepatol 2021; 75:1128–1141.

34. van de Veen W, Globinska A, Jansen K, Straumann A, Kubo T, Verschoor D, Wirz OF, Castro-Giner F, Tan G, Rückert B, Ochsner U, Herrmann M, Stanić B, van Splunter M, Huntjens D, Wallimann A, Fonseca Guevara RJ, Spits H, Ignatova D, Chang YT, Fassnacht C, Guenova E, Flatz L, Akdis CA, Akdis M. A novel proangiogenic B cell subset is increased in cancer and chronic inflammation. Sci Adv 2020;6, eaaz3559.

35. Vanuytsel T, Senger S, Fasano A, Shea-Donohue T. Major signaling pathways in intestinal stem cells. Biochim Biophys Acta 2013;1830:2410–2426.

36. He Z, Chen L, Chen G, Smaldini P, Bongers G, Catalan-Dibene J, Furtado GC, Lira SA. Interleukin 1 beta and matrix metallopeptidase 3 contribute to development of epidermal growth factor receptor-dependent serrated polyps in mouse cecum. Gastroenterology 2019; 157:1572–1583.e8.

37. Wang Z, Xiong S, Mao Y, Chen M, Ma X, Zhou X, Ma Z, Liu F, Huang Z, Luo Q, Ouyang G. Periostin promotes immunosuppressive premetastatic niche formation to facilitate breast tumour metastasis. J Pathol 2016; 239:484–495.

38. Nakanishi Y, Diaz-Meco MT, Moscat J. Serrated colorectal cancer: the road less travelled? Trends Cancer 2019;5:742–754.

39. Lee HO, Hong Y, Etlioglu HE, Cho YB, Pomella V, Van den Bosch B, Vanhecke J, Verbandt S, Hong H, Min JW, Kim N, Eum HH, Qian J, Boeckx B, Lambrechts D, Tsantoulis P, De Hertogh G, Chung W, Lee T, An M, Shin HT, Joung JG, Jung MH, Ko G, Wirapati P, Kim SH, Kim HC, Yun SH, Tan IBH, Ranjan B, Lee WY, Kim TY, Choi JK, Kim YJ, Prabhakar S, Tejpar S, Park WY. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. Nat Genet 2020;52:594–603.

40. Li X, Sun X, Carmeliet P. Hallmarks of endothelial cell metabolism in health and disease. Cell Metab 2019; 30:414–433.

41. Perillo B, Di Donato M, Pezone A, Di Zazzo E, Giovannelli P, Galasso G, Castoria G, Migliaccio A. ROS in cancer therapy: the bright side of the moon. Exp Mol Med 2020;52:192–203.

42. Sheu JJ, Guan B, Tsai FJ, Hsiao EY, Chen CM, Seruca R, Wang TL, Shih IM. Mutant BRAF induces DNA strand breaks, activates DNA damage response pathway, and up-regulates glucose transporter-1 in nontransformed epithelial cells. Am J Pathol 2012;180:1179–1188.

43. Milkovic L, Zarkovic N, Saso L. Controversy about pharmacological modulation of Nrf2 for cancer therapy. Redox Biol 2017;12:727–732.

44. Fukuyo Y, Inoue M, Nakajima T, Higashikubo R, Horikoshi NT, Hunt C, Usheva A, Freeman ML, Horikoshi N. Oxidative stress plays a critical role in inactivating mutant BRAF by geldanamycin derivatives. Cancer Res 2008;68:6324–6330.

45. Burgener SS, Leborgne NGF, Snipas SJ, Salvesen GS, Bird PI, Benarafa C. Cathepsin G inhibition by Serpinb1 and Serpinb6 prevents programmed necrosis in neutrophils and monocytes and reduces GSDMD-driven inflammation. Cell Rep 2019;27:3646–3656.e5.

46. Gehart H, Clevers H. Tales from the crypt: new insights into intestinal stem cells. Nat Rev Gastroenterol Hepatol 2019;16:19–34.

47. Jackstadt R, van Hooff SR, Leach JD, Cortes-Lavaud X, Lohuis JO, Ridgway RA, Wouters VM, Roper J, Kendall TJ, Roxburgh CS, Horgan PG, Nixon C, Nourse C, Gunzer M, Clark W, Hedley A, Yilmaz OH, Rashid M, Bailey P, Biankin AV, Campbell AD, Adams DJ, Barry ST, Steele CW, Medema JP, Sansom OJ. Epithelial NOTCH signaling rewires the tumor microenvironment of colorectal cancer to drive poor-prognosis subtypes and metastasis. Cancer Cell 2019;36:319–336.e7.

48. Kane AM, Liu C, Fennell LJ, McKeone DM, Bond CE, Pollock PM, Young G, Leggett BA, Whitehall VLJ. Aspirin

reduces the incidence of metastasis in a pre-clinical study of Braf mutant serrated colorectal neoplasia. Br J Cancer 2021;124:1820–1827.

49. Kim JH, Kim K-J, Rhee Y-Y, Bae JM, Cho N-Y, Lee HS, Kang GH. Gastric-type expression signature in serrated pathway-associated colorectal tumors. Hum Pathol 2015;46:643–656.

50. Yu S, Balasubramanian I, Laubitz D, Tong K, Bandyopadhyay S, Lin X, Flores J, Singh R, Liu Y, Macazana C, Zhao Y, Béguet-Crespel F, Patil K, Midura-Kiela MT, Wang D, Yap GS, Ferraris RP, Wei Z, Bonder EM, Häggblom MM, Zhang L, Douard V, Verzi MP, Cadwell K, Kiela PR, Gao N. Paneth cell-derived lysozyme defines the composition of mucolytic microbiota and the inflammatory tone of the intestine. Immunity 2020;53:398–416.e8.

51. Morikawa R, Nemoto Y, Yonemoto Y, Tanaka S, Takei Y, Oshima S, Nagaishi T, Tsuchiya K, Nozaki K, Mizutani T, Nakamura T, Watanabe M, Okamoto R. Intraepithelial lymphocytes suppress intestinal tumor growth by cell-to-cell contact via CD103/E-cadherin signal. Cell Mol Gastroenterol Hepatol 2021; 11:1483–1503.

52. Ayukawa S, Kamoshita N, Nakayama J, Teramoto R, Pishesha N, Ohba K, Sato N, Kozawa K, Abe H, Semba K, Goda N, Fujita Y, Maruyama T. Epithelial cells remove precancerous cells by cell competition via MHC class I–LILRB3 interaction. Nat Immunol 2021; 22:1391–1402.

53. Chen B, Scurrah CR, McKinley ET, Simmons AJ, Ramirez-Solano MA, Zhu X, Markham NO, Heiser CN, Vega PN, Rolong A, Kim H, Sheng Q, Drewes JL, Zhou Y, Southard-Smith AN, Xu Y, Ro J, Jones AL, Revetta F, Berry LD, Niitsu H, Islam M, Pelka K, Hofree M, Chen JH, Sarkizova S, Ng K, Giannakis M, Boland GM, Aguirre AJ, Anderson AC, Rozenblatt-Rosen O, Regev A, Hacohen N, Kawasaki K, Sato T, Goettel JA, Grady WM, Zheng W, Washington MK, Cai Q, Sears CL, Goldenring JR, Franklin JL, Su T, Huh WJ, Vandekar S, Roland JT, Liu Q, Coffey RJ, Shrubsole MJ, Lau KS. Differential pre-malignant programs and microenvironment chart distinct paths to malignancy in human colorectal polyps. Cell 2021;184:6262–6280.e26.

54. Hu X, Li YQ, Li QG, Ma YL, Peng JJ, Cai SJ. ITGAE defines CD8$^+$ tumor-infiltrating lymphocytes predicting a better prognostic survival in colorectal cancer. EBioMedicine 2018;35:178–188.

55. Toh JWT, Ferguson AL, Spring KJ, Mahajan H, Palendira U. Cytotoxic CD8$^+$ T cells and tissue resident memory cells in colorectal cancer based on microsatellite instability and BRAF status. World J Clin Oncol 2021; 12:238–248.

56. Okła K, Farber DL, Zou W. Tissue-resident memory T cells in tumor immunity and immunotherapy. J Exp Med 2021;218:e20201605.

57. Pecci F, Cantini L, Bittoni A, Lenci E, Lupi A, Crocetti S, Giglio E, Giampieri R, Berardi R. Beyond microsatellite instability: evolving strategies integrating immunotherapy for microsatellite stable colorectal cancer. Curr Treat Options Oncol 2021;22:69.

58. Corcoran R, Giannakis M, Allen J, Chen J, Pelka K, Chao S, Meyerhardt J, Enzinger A, Enzinger P, McCleary N, Yugelun M, Abrams T, Kanter K, Van Seventer E, Bradford W, Fetter I, Siravegna G, Tian J, Clark J, Ryan D, Hacohen N, Parikh A. SO-26 Clinical efficacy of combined BRAF, MEK, and PD-1 inhibition in BRAFV600E colorectal cancer patients. Ann Oncol 2020; 31:S226–S227.

59. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, Angell H, Fredriksen T, Lafontaine L, Berger A, Bruneval P, Fridman WH, Becker C, Pages F, Speicher MR, Trajanoski Z, Galon J. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 2013;39:782–795.

60. Yano H, Andrews LP, Workman CJ, Vignali DAA. Intra-tumoral regulatory T cells: markers, subsets and their impact on anti-tumor immunity. Immunology 2019; 157:232–247.

61. Bettington M, Walker N, Clouston A, Brown I, Leggett B, Whitehall V. The serrated pathway to colorectal carcinoma: current concepts and challenges. Histopathology 2013;62:367–386.

62. WHO Classification of Tumours Editorial Board. WHO classification of tumours: digestive system tumours. Lyon: International Agency for Research on Cancer, 2019.

63. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, Fallahi-Sichani M, Dutton-Regester K, Lin JR, Cohen O, Shah P, Lu D, Genshaft AS, Hughes TK, Ziegler CG, Kazer SW, Gaillard A, Kolb KE, Villani AC, Johannessen CM, Andreev AY, Van Allen EM, Bertagnolli M, Sorger PK, Sullivan RJ, Flaherty KT, Frederick DT, Jané-Valbuena J, Yoon CH, Rozenblatt-Rosen O, Shalek AK, Regev A, Garraway LA. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 2016;352:189–196.

64. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16:284–287.

65. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013;14:7.

66. Hong J, Guo F, Lu SY, Shen C, Ma D, Zhang X, Xie Y, Yan T, Yu T, Sun T, Qian Y, Zhong M, Chen J, Peng Y, Wang C, Zhou X, Liu J, Liu Q, Ma X, Chen YX, Chen H, Fang JYF. nucleatum targets lncRNA ENO1-IT1 to promote glycolysis and oncogenesis in colorectal cancer. Gut 2021;70:2123–2137.

67. You Z, Li Y, Wang Q, Zhao Z, Li Y, Qian Q, Li B, Zhang J, Huang B, Liang J, Chen R, Lyu Z, Chen Y, Lian M, Xiao X, Miao Q, Fang J, Lian Z, Eric Gershwin M, Tang R, Ma X. The clinical significance of hepatic CD69(+) CD103(+) CD8(+) resident-memory T cells in autoimmune hepatitis. Hepatology 2021;74:847–863.

68. Chen YP, Yin JH, Li WF, Li HJ, Chen DP, Zhang CJ, Lv JW, Wang YQ, Li XM, Li JY, Zhang PP, Li YQ, He QM, Yang XJ, Lei Y, Tang LL, Zhou GQ, Mao YP, Wei C, Xiong KX, Zhang HB, Zhu SD, Hou Y, Sun Y, Dean M, Amit I, Wu K, Kuang DM, Li GB, Liu N, Ma J. Single-cell transcriptomics reveals regulators underlying immune cell diversity and

immune subtypes associated with prognosis in nasopharyngeal carcinoma. Cell Res 2020;30:1024–1042.