Comprehensive dissection of immune microenvironment in the progression of early gastric cancer at spatial and single-cell resolution

Tiantian Du^{#1}, Huiru Gao^{#1}, Honglei Wu^{#4}, Juan Li¹, Peilong Li¹, Jie Gao¹, Qiuchen Qi¹, Xiaoyan Liu¹, Lutao Du^{*1,2,3}, Yunshan Wang^{*1}, Chuanxin Wang^{*1,2,3}

¹ *Department of Clinical Laboratory, The Second Hospital, Cheeloo College of Medicine, Shandong University, Jinan, Shandong, 250033, China.*

² *Shandong Engineering & Technology Research Center for Tumor Marker Detection, 250033, Jinan, Shandong, China.*

³ *Shandong Provincial Clinical Medicine Research Center for Clinical Laboratory, 250033, Jinan, Shandong, China.*

⁴ *Department of Gastroenterology, The Second Hospital of Shandong University, 250033, Jinan, Shandong, China.*

[#]These authors contributed equally to this work.

^{*}**Corresponding Author:** Chuanxin Wang, Dr., Pro.

Department of Clinical Laboratory, The Second Hospital, Cheeloo College of Medicine, Shandong University, 247 Beiyuan Street, Jinan 250033, Shandong, China.

[Tel:+86-0531-85875002](tel:+86-0531-85875002), Fax: +86-0531-88962544, E-mail: cxwang@sdu.edu.cn

^{*}**Corresponding Author:** Yunshan Wang, Pro.

Department of Clinical Laboratory, The Second Hospital, Cheeloo College of

Medicine, Shandong University, 247 Beiyuan Street, Jinan 250033, Shandong, China.

Tel: +86-17660085123, Fax: +86-17660085123, E-mail: wangyunshan135@126.com

*

Corresponding Author: Lutao Du, Dr., Pro.

Department of Clinical Laboratory, The Second Hospital, Cheeloo College of

Medicine, Shandong University, 247 Beiyuan Street, Jinan 250033, Shandong, China.

Tel: +86-17660085177, Fax: +86-17660085177, E-mail: lutaodu@sdu.edu.cn

Abstract

While the changes of tumor immune microenvironment (TME) have critical implications for most tumor progression, works that could reveal the compositions and immunity features of TME are needed. Profiling gastric malignant cells at single-cells resolution has shown the transcriptional heterogeneity is represented at different states of gastric cancer, implying that diverse cell states may exist, including immune cells, and all components in TME make some balances in early gastric cancer (EGC) progression. However, it remains unclear how immune cells contributing malignancy of gastritis, constituting general characteristics of gastric TME. Furthermore, the role of specific interactions among cells in gastric TME remains to be illustrated. Here, we performed spatial transcriptomes and single-cell RNA-seq analysis across 18 gastric samples, identifying 17 celltypes and reconstructing their location information. We found that immune cells represented different degree of dysregulations during the progression from non-atrophic gastritis (NAG), atrophic gastritis (AG) to EGC, including imbalance of cytotoxic and inhibitory effects in T cells, maturation inhibition in B cells and malignant genes up-regulated obviously in myeloid cells. Besides, pathway activities showed that hypoxia, reactive oxygen species and fatty metabolism signaling were activated from AG stage, which may accelerate progression of EGC. Moreover, cellular interactions further identified the roles of hypoxia in gastric TME. Overall, the multi-omics data presented in this study offer a comprehensive view of immune cell types, states changes and locations within the gastric tissues during the progression from NAG, AG to EGC, advancing our understanding of the composition and immunity of different gastric states, offering diagnostic and preventive thoughts for EGC.

Key words: Human gastric cancer; Tumor progression; Immune microenvironment; Single-cell RNA-seq; Visium spatial transcriptomics

Introduction

Gastric cancer possesses the characteristic of high mortality, is the fifth most common cancer. The development of gastric cancer is a multi-step process, generally from non-atrophic gastritis (NAG) to atrophic gastritis (AG) with or without intestinal metaplasia (IM), and finally to all-stage gastric cancer (GC), which is academically known as Correa cascade ¹. The early gastric cancer (EGC), as initial phase of the process of GC, holds well prognosis more than 90% ^{2 3}. Thus, a comprehensive understanding of EGC progression from NAG, on the one hand, is the base of further timely intervention before advanced stages, hopefully improving GC condition ⁴. On the other hand, it is in favor of exploitation of targets for GC risk warning and noninvasive diagnosis, as the most recognized means of EGC diagnosis now remains the invasive endoscopic biopsy ⁴⁵.

The progression of most tumors is known as results of dynamic interaction between tumor cells and their immune microenvironment, including EGC ⁶. During the process, gastric mucosa tissues appear heterogeneous changes and malignant cells, which have been reported previously ⁷⁻¹¹. However, how immune microenvironment play roles during the EGC progression is still unclear. It was showed that infiltration of inflammatory cells, as well as expression of chemokines and interleukins are different among normal gastric mucosa, simple gastritis, AG and GC ¹². IL33 was reported to promote metaplasia and its exogenous expression could moderately induce spasmodic polypeptide-expressing metaplasia (SPEM), may represent as important joint of chronic gastritis and IM ¹³. The typically binary M1/M2 paradigm of macrophages was disappeared, and transcriptionally heterogeneous among each cluster was in its place ^{14 15}. Besides, 12-gene prognostic signature for metastatic gastric adenocarcinoma (GAC) was established ¹⁶. Although few works have been reported, they are not enough to illustrate changes during progression from NAG to EGC.

As the complexities of gastric immune microenvironment and technical limitations, previously, it was difficult to profile the component, structure and immunity of immune microenvironment. The recently developed single-cell

sequencing (scRNA-seq), VDJ and spatial transcriptomics (ST) sequencing have realized the identity of complicated immune microenvironment. As new techniques, scRNA-seq requires digestive cell suspension and lacks spatially relative location, while ST sequencing represents with spatial spots images with low cell resolution. Then, with the comprehensive use from the dimension of multi-omics, not only the cell biology is elucidated accurately at single cell resolution with scRNA-seq, but the spatial location information and intercellular interaction is acquired by use of ST, even the immunized libraries of T/B cells are obtained ^{16 17}. Several tumor and immune microenvironment profiles have been successfully clarified by combined scRNA-seq and ST, including pancreatic cancer ¹⁸, squamous cell carcinoma ¹⁹, as well as colorectal cancer liver metastasis ²⁰. Thus, we considered application of the combined methods to elucidate microenvironment characteristics during development from NAG, AG to EGC.

In this study, we acquired 21,848 single cells and 11,347 spatial spots. With the combined use of scRNA-seq and ST sequencing, we characterized the dysregulation of gastric immune microenvironment changes. The multi-omics data offer a comprehensive view of immune cell types, states change and locations within the gastric tissues during the progression from NAG, AG to EGC, advancing our understanding of the composition and immunity of different gastric states, offering diagnostic and preventive thoughts for EGC.

Methods

Experimental models and subject details

Human sample handling (samples and sample information)

Clinical samples were received immediately after extraction from the endoscopy (The delivery time of every sample was less than 30 min) and processed rapidly to ensure viability and quality for scRNA-seq and ST ²¹.

The 12 tissue samples used for ST in our study were randomly selected from

established cohort more than 200 specimens, and other 6 fresh tissues used for scRNA-seq were randomly collected from the clinic. All study participants provided written informed consent and study protocols were approved by the ethical review community of the Second Hospital of Shandong University. All of freshly resected biopsy specimens were obtained from one or two sites by conventional upper gastrointestinal endoscopy. Fresh samples were then washed several times with dulbecco's phosphate-buffered saline (DPBS) for further embedding or digestion ⁷.

Pathological annotation for spatial images

All sections used in the spatial study were identified by two independent pathologists. Tissue HE images included tumor cells characteristics were defined as the EGC group, no tumor features but contained intestinal metaplasia areas were defined as AG group, and tissues with mild inflammatory but no tumor and intestinal metaplasia areas were defined as NAG group.

Spatial Transcriptomics

Slide preparation

For 12 samples used for the ST, trimming them to a suitable size, then embedded with Optimal Cutting Temperature (OCT) embedding solution (Sakura, #4532, USA) and transfer to -80 °C refrigerator for quick freezing and long-term storage. Cryosections of the above tissues were cut at 10 μm thickness, 10 consecutive slices for RNA extraction to assess the RNA integrity number (RIN) ($RIN \geq 7$ is qualified), and the other 2 consecutive slices for hematoxylin-eosin staining (H&E) staining to confirm the detailed structures of these tissues.

The slides in the Visum spatial gene expression slide and Reagent Kit (10x Genomics, PN-1000184) contained four identical 6.5×6.5 mm tissue capture areas, each with ~5000 spots measuring a diameter of 55 μm and the distance between two spots is 100 μm, which should be expected to encompass 6-10 total cells. An individual spot containing millions of oligonucleotides with features: a partial Illumina TruSeq Read 1 sequence, 16 nt Spatial Barcode (shared within each specific spot), 12 nt unique molecular identifier (UMI) (Identification of duplicate molecules), and 30 nt poly (dT) sequence (captures polyadenylated mRNA for cDNA synthesis).

Optimization of the permeabilization time

An optimization slide totally covers eight capture areas, the same size as the GEX slide. six capture areas for cryosections cut at 10 μ M thickness, the other two are used for positive and negative controls. An optimal permeabilization time must be identified before a new tissue used for generating Visium Spatial Gene Expression libraries. Briefly, the Visium Spatial Tissue Optimization workflow described as placing tissue sections on six capture areas on a Visium Tissue Optimization slide (from Visium Spatial Gene Expression Reagent Kit, 10x Genomics, PN-1000186). The prepared sections were further fixed, stained, and then permeabilized for different times. The time that resulted in maximum fluorescence signal with the lowest signal diffusion was used as the best permeabilization time, the longer time will be optimal if the signal was same at multiple time points. There were different time points among these tissues, just listed as following: for N22-0, N22-1, N8-1, N21-1, N34-1, C3, C7, C8, C9, 24 min was used; while for N26-0, N28-0, N24-1, 22 min was used.

Spatial staining and imaging

Sectioned slides were incubated at 37°C with the active surface facing up for 1 min, and fixed with precooled methyl alcohol in -20°C for 30 min, followed by H&E staining (Eosin, Dako CS701, Hematoxylin Dako S3309, bluing buffer CS702), sections were processed at RT by isopropanol (MilliporeSigma) for 1 min, by hematoxylin (Agilent, Dako S3309) for 7 min, by Bluing Buffer (Agilent, CS702) for 2 min, and by Eosin Mix (MilliporeSigma) for 1 min. Last, slides were incubated for 5 min at 37°C in the Thermocycler Adaptor (10x Genomics, PN-3000380). RNase and DNase free water were used in every washing step. After air-drying, the bright-field images were taken on a Leica DMI8 whole-slide scanner (Germany) at 10 $\times$ resolution.

Permeabilization and reverse transcription

The slides were inserted into slide cassettes (from the from the Visium Slide Kit) to ensure individual reaction chambers of each tissue section. Then according to the predetermined optimal permeabilization time, the tissue sections on the slides were subjected to 70 μ L permeabilization enzyme (from the Visium Reagent Kit) at 37°C

for different times to capture polyadenylated mRNA by the primers on the spots. After incubation, each well was washed with 100 μ L 0.1 $\times$ SSC (Sigma-Aldrich), and added with 75 μ L reverse transcription Master Mix (from the Visium Reagent Kit) to generate spatially barcoded cDNA.

cDNA library preparation and sequencing

After the RT Master Mix was removed from the wells at the end of first-strand synthesis, each well was incubated in 75 μ L 0.08 M KOH for 5 min and washed with 100 μ L EB buffer (QIAGEN) at room temperature (RT). Then 75 μ L Second Strand Mix was added to each well for second-strand synthesis. cDNA amplification and library construction were performed on a S1000TM Touch Thermal Cycler (Bio Rad). Once the Visium Spatial Gene Expression libraries were constructed using Visum spatial Library construction kit (10x Genomics, PN-1000184), further spatial transcriptome sequencing was processed on the Illumina Novaseq6000 platform immediately. The standard paired-end sequencing reached a depth of at least 100,000 reads per spot (pair-end 150 bp, PE150) (performed by CapitalBio Technology, Beijing).

Processing of the raw ST data

Raw Visium Spatial RNA-seq data, bright-field and fluorescence microscope images were processed by the Space Ranger software (version 6.1, 10x Genomics) to align reads, filter and construct feature-spot matrixes, UMIs and genes counting, and match spots on the spatial slide images. The GRCh38 v86 genome assembly was used as reference. The QCs were applied before further clustering and gene expression were analyzed by the Seurat R package (version 4.0). Briefly, the spots away from the main section region were discarded. The mitochondrial and ribosomal genes and genes were expressed ≤ 3 spots were filtered. After quantity control, we obtained 11,347 spots for further analysis.

scRNA-seq

Human Gastric tissues dissociation protocol

All gastric tissues were distinguished from non-tumor and tumor regions by gastroscopist before further processing. After several washing with 1 $\times$ DPBS

containing 0.04% bovine serum albumin (BSA, Miltenyi Biotec, 130-091-376), gastric tumor tissues single cells were acquired with digestive enzymes containing 200 μ L Enzyme H, 100 μ L Enzyme R, 25 μ L Enzyme A (all provided in the human tumor dissociation kit (Miltenyi Biotec; 130-095-929), and 4.7 mL DMEM (Gibco; 8117133). While non-tumor tissues digestion was performed in mixed enzymes containing 100 μ L Enzyme D, 50 μ L Enzyme R, 12.5 μ L Enzyme A (all provided in the human multi-tissues dissociation kit (Miltenyi Biotec; 130-110-201), and 2.35 mL DMEM (Gibco; 8117133). Single cells were generated from shaking incubation in the water bath at 37 °C. After the digestive reaction was deactivated by adding another 45 mL fresh DMEM, the cell suspensions were filtered using a 70 μ m cells strainer (Miltenyi Biotec; 130-098-462) and then centrifuged at 300g for 5 min. Live nuclear cells were enriched using a Dead Cell Removal kit (Miltenyi Biotec; 130-090-101) and Red Cell Removal kit (Solarbio; R1010). The final cells were re-suspended in 1 $\times$ DPBS containing 0.04% BSA and prepared for further steps.

Library preparation and droplet based scRNA-seq sequencing

Briefly, the droplet based sequencing platform was used in these studies. Gastric tissue suspensions were loaded on a Chromium Single Cell Controller instrument (10x Genomics, Pleasanton, California, USA) to generate single-cell gel beads in emulsion (GEMs). GEM-RT was performed on a S1000TM Touch Thermal Cycler (Bio Rad): 53°C for 45 min, 85°C for 5 min; held at 4°C and stored at -20°C. Then the GEMs were broken and the barcoded, full-length cDNA was amplified, and quality assessed using an Agilent 4200 (CapitalBio Technology, Beijing). scRNA-Seq libraries were prepared using the Chromium Single Cell 5' Library & Gel Bead Kit (P/N 1000006, 10x Genomics), Chromium Single Cell A Chip Kit (10x Genomics, 120236), Human T Cell (1000005) and Single Cell V(D)J Enrichment Kit, and Human B Cell (1000016) and Single Cell V(D)J Enrichment Kit according to manufacturer's instructions. The libraries were finally sequenced using an Illumina Novaseq6000 platform with a sequencing depth of at least 100,000 reads per cell with PE150 reading strategy (performed by CapitalBio Technology, Beijing).

Processing of raw scRNA-seq data

Single-cell gene expression quantification, quality control

Alignment, filtering, barcode counting, and unique molecular identifier (UMI) counting of the sequencing data were performed with Cell Ranger software package and mapped to the hg38 reference genome to generate feature-barcode matrix at the base of human reference genome (GRCh38). Then secondary clustering and differential expression analysis of the raw gene expression matrices were further processed with the Seurat R package. Cells with either gene number ranked over the top 1% or fewer than 200, or mitochondrial gene ratio was more than 25% were regarded as unqualified and excluded before downstream analysis. After quantity control, we obtained 28,860 high-quality single cells for further analysis.

Integration of multiple samples, batch correction and dimension reduction

We separately used SCTransform (R package, version 0.3.2) and harmony (R package, version 0.1.0) to integrate the ST data and scRNA-seq data from different patients. For scRNA-seq data, the new combined feature-barcode matrix was first applied for normalization, FindVariableFeatures (12000 highly variable genes were identified) and scaling. For spatial data, “SCTransform” function replaced the above procession. Next, principal component analysis (PCA) (RunPCA, from variable genes) was performed. The top 30 principal components (PCs) were selected for FindNeighbors, FindClusters, UMAP (RunUMAP) and TSNE (RunTSNE) to obtain clusters and bidimensional coordinates for each spot or cell and visualization. In this study, the sample source was used defined as batch factor and using “RunHarmony” or “SCTransform” function to reduce the impact of batch effect and generate batch corrected data for downstream analysis.

Cell type annotation

Cell type was annotated by referring to published papers manually and singleR (version 1.6.1). For singleR, unbiased cell type recognition from scRNA-seq data was performed by leveraging reference transcriptomic datasets. The known markers used for cell cluster annotation are listed in Supplementary information.

BCR and TCR analysis

With the 10 x Genomics 5' V(D)J platform, the immune libraries were built and

TCR/BCR sequences were assembled following Cell Ranger VDJ protocol. Assembled contigs were filtered by intersected with barcodes in scRNA-seq. The retained cells were that with at least one TCR α chain (TRA) and one TCR β chain (TRB). For a given T cell, if there were two or more α or β chains assembled, the highest expression level α or β chains was regarded as the dominated α or β chain in the cell. Each unique α - β pairs was defined as a clonotype. T cells with exactly same clonotype were composed a T cell clone.

Similarly, as for BCR clonotype, only cells with at least one heavy chain (IGH) and one light chain (IGK/IGL) were remained. For a given B cell, if there were two or more IGH or IGK/IGL assembled, the highest expression level IGH or IGK/IGL was defined as the dominated IGH or IGK/IGL in the cell. Each unique IGH-IGK/IGL pair was defined as a clonotype. B cells with exactly same clonotype composed a B cell clone.

Marker gene detection and differential expression analyses

The FindMarkers function with default parameters ($\log_{fc}.\text{threshold} = 0.01$) was used to identified different expression genes (DEGs) between two idents. And FindAllMarkers function (wilcox method and $\log_{fc}.\text{threshold} = 0.25$ were used) and was called to identified marker genes of each cluster. Genes with adjusted $P < 0.05$ as the significant ones and reserved.

Pathway enrichment analysis

The pathway activities of cells were quantified by GSEA, GSVA and AUCell package, respectively. The log-transformed normalization expression matrixes of cells or spots were as inputs and all analysis were performed with default parameters. The referenced gene sets, including 50 hallmark signatures (MSigDB, H sets), 186 KEGG signatures (MSigDB, C2 sets), 1599 Reactome pathways (MSigDB, C2 sets), and 10402 GO signatures (MSigDB, C5, GO sets), were all derived from the Molecular Signatures Database (MSigDB v7.1).

Trajectory analysis

The pseudotime trajectories were processed both with Monocle (R package, Version 2.20.0) and Monocle 3 (R package, Version 1.0.1) to determine different cell states.

The counts data of gene-barcode matrix were used as input. For Monocle, the function of `newCellDataSet` was used to create a new data object for analysis, then two components and “DDRTree” were used to reduce dimensions, followed by “orderCells”, the results of cell developmental trajectory were presented. The `lowerDetectionLimit` is 0.5. While for Monocle 3, the `new_cell_data_set` function was used for objects building, the starting point of the developmental trajectory based on the biological background.

Cell-cell interaction analysis

cell-cell communication atlases, which contained pairs of ligands and receptors between different cells, were built and visualized by CellChat (R package, Version 1.1.3)²². The ligand-receptor interactions included three modules named secreted signaling, ECM receptor signaling, and cell-cell touch signaling, which consisted of multimeric ligand-receptor complexes, soluble agonists and antagonists, as well as stimulatory and inhibitory membrane-bound co-receptors. The min cell number is 3 in `filterCommunication` function.

Transcription factor module analysis

R packages including “SCENIC”, “RcisTarget”, “GENIE3” and “AUCell” were used to analyze the transcription factor of scRNA-seq and ST data. The reference data named “hg38__refseq-r80__10kb_up_and_down_tss.mc9nr.feather” and “hg38__refseq-r80__500bp_up_and_100bp_down_tss.mc9nr.feather” which downloaded from (<https://resources.aertslab.org/cistarget/>) were used. In the following analysis, we filtered genes, computed a gene-gene correlation matrix, performed transcription factor network analysis and calculated a score for each TF module in each individual cell.

inferCNV analysis

The CNVs of each cell and each spot were estimated by “inferCNV” package at the basis of gene transcriptome profiles²³. For whole cells and spots among three groups (NAG, AG, EGC), the NAG groups were defined as references. Then, for evaluating of somatic hypermutation (SHM), the B cells were defined as controls.

Spatial Transcriptomics Spots Deconvolution using SPOTlight

SPOTlight was a R package that deconvoluted spots in ST data using scRNA-seq data by Seeded NMF regression ²⁴. First, we acquired the marker genes of each cell type in scRNA-seq which was used to deconvoluted the spots in the following by the function called “spotlight_deconvolution”. After adding the percent of cell type in each spot to the metadata of Seurat object, we employed the “SpatialFeaturePlot” to present the cell proportion of ST.

Correlation analysis

Correlation analysis for different regions, genes or gene sets were all operated by ‘Spearman’ methods of the stat_cor function.

Multiplex Immunofluorescence

The embedded patient tissues selected randomly from those subjected to spatial transcriptomes sequencing and biobank were sectioned to 6-μm-thick slides. Then the sections were dried for 10 min at room temperature (RT), circled and soaked for 15 min in PBS to removing OCT. After washing 3 times in PBS, the sections were blocked with 10% BSA/PBS for 1 hour at RT with or without washing. The well-blocked sections were then incubated overnight at 4 °C with the first primary antibodies. Next day, after 3 times washing, the sectioned were incubated with the corresponding second antibodies for 2 hours. After 3 times shaker washing, the sectioned were incubated with the second primary antibodies and corresponding second antibodies just repeated steps for the first primary. Lastly, DAPI was used to stained nuclei for 8 min at RT and immunofluorescence images were obtained using digital slice scanners. The primary antibodies used in the assay were listed as following: Anti-Syndecan-1 (1:200, Abcam, ab128936), Anti-pan Cytokeratin (1:200, Abcam, ab7753), and DAPI (Servicebio, G1012).

Data and code availability

Processed data can be accessed from the NCBI Gene Expression Omnibus (GEO) database when the study acquires permission. The major codes generated during the process will be provided if needed.

Results

Exploring the spatial architecture and single cell characteristics of the immune microenvironment in the progression of EGC

To explore the mechanism of transformation from gastritis to EGC, we identified the spatial features of 12 tissue samples from NAG, AG, and EGC patients, separately, which represented different stages in the development of gastric cancer (Figure 1A-C, Supplementary Figure 4A). We recognized distinguishable tumor regions, atrophic regions with or without IM, and regions with no obvious lesions (Figure 1F). Correlation analysis of these regions showed distinctions among the non-lesion regions of each groups much less the distinguishable lesions areas (Supplementary Figure 5A). Then, globally, pathway activities profiles of the spatial regions showed specific activated patterns (Supplementary Figure 1D-E). Comparing with NAG groups, in EGC groups, the MYC signaling, oxidative phosphorylation, interferon alpha response, IL-1 signaling, mTORC1 signaling, TNF signaling, and myeloid leukocyte mediated immunity were up-regulated, while in AG, fatty acid metabolism, xenobiotic metabolism, and lipids metabolism were up-regulated. Inversely, the pathway of KRAS signaling down, aspartic type peptidase activity, B cell receptor signaling, and cell recognition, which up-regulated in NAG groups were all down-regulated in EGC groups (Supplementary Figure 1D-E). Although we found the pathway activities were partly associated with specific celltypes, such as the muscular regions were in line with activated coagulation pathway. For most pathways, their activities were inhomogeneous in each spot and had difficult in matching with particular celltypes, on the one hand for the technical limitations, on the other hand for the complex microenvironment. Besides, there was an evolutionary trend from NAG, AG, to EGC cells analyzed by time trajectory analysis, along with distinct changes of gene expression (Figure 1D-E). Given that, we further explored the difference of cell characteristics and gene expression profiles among three groups with scRNA-seq. We harvested 28,860 cells using 10x scRNA-seq from 6 fresh gastroscopic biopsy specimens of NAG, AG and EGC (Figure 1A). Then we merged,

filtered, normalized, batch-corrected, clustered all cells, finally identified 17 cell types of 21,848 cells that passed the quality control (Figure 2A). Epithelium cells, T cells, plasmas occupy a larger portion, fibroblasts and parietal cells have similar number of cells. Besides, there exist endothelium, B cells, mast cells, goblet cells, chief cells and so on, which in a relatively small number (Figure 2E). The distribution of each cell type in different groups and samples was even, indicating no batch effect in the scRNA-seq data (Figure 2B-C). The expression of signature genes shows the correctness of annotation (Figure 2D).

We explored the origin of each celltype and found considerable heterogeneity between the different groups (Figure 2E). Epithelium cells occupied the largest number in gastric tissues, and most of them were from EGC. Existing in gastric glands, parietal cells were consistent with epithelium cells in specimen sources. For fibroblasts, endothelium cells, and pericytes, there was a decrease from NAG, AG, to EGC, indicating a gradual reduction in vascular and smooth muscle tissues, which may due to sampling differences for biopsy specimens were slightly tiny. For chief cells, the reduction was tally with the signature of SPEM. Furthermore, we explored the expression of TFF2 (Supplementary Figure 6D) which referred in GC progression of SPEM which means SPEM may occurred in our research. In addition, it was worth noting that a group called proliferative cells with the highest proportion in EGC samples, indicating increased ability of proliferation in the development of gastric cancer. Meanwhile, proliferative cells significantly expressed T cell markers such as CD3D and CD3E (Figure 2D), which implied T cell was stimulated in the progression of gastric cancer in proliferation ability.

Then, using SPOTlight, we deconvoluted the ST data combining scRNA-seq data. As a result, we acquired celltype components of each spot. Epithelium cells were widely distributed in spots except for the muscular layer and lymphocyte aggregation area. T cells and B cells were located near each other, and we observed that in the lymphocyte aggregation area, T cells and B cells occupied a relatively high proportion. Compared with the T cells and B cells, plasmas existed in the outer position in the lymphocyte aggregation area. Smooth muscle cells existed mainly in the muscle layer

which was reflected in the HE, as it proved the accuracy of cell localization to some extent. The co-localization of Schwann cells, pericytes and fibroblasts indicated that these cells and smooth muscle may be the main cell components of muscle layer. Macrophages tended to exist more at the edge of tissue compared with other immune cells, and the chief cells also avoided the area of lymphocyte aggregation, which was consistent with the epithelium cells. Besides, we also noticed that in NAG, proliferative cells mainly gathered in the lymphocyte aggregation region, but gathered in the epithelial region in the other two groups, suggesting proliferative cells may have different phenotypes and functions including immune cell phenotypes in NAG and Epithelium cell phenotypes in AG and EGC. There also existed interesting phenomenon that the lymphocyte aggregation area was largest in NAG, but diminished in AG and EGC, indicating the immune defense function was the strongest in NAG (Figure 2F and Supplementary Figure 6F).

Previous studies had referred that tumors progression were often accompanied with increased copy number variations (CNV). However, the CNVs had few effects on GC progression, partly given the globally low variations in gastric tissues. In line with the previous work, there were no significant difference of CNV scores during the progression from NAG to EGC (Supplementary Figure 4B-E). The data indicated that there were obvious changes in gene expression profiles and cell characteristics of microenvironment among NAG, AG, and EGC groups.

Different expression patterns of T cells from NAG to EGC

We noticed that T cells, which play crucial role in tumor microenvironment, accounted for the largest proportion of immune cells. Then, we extracted all the T cells from three groups then reduced the dimensionality and re-clustered them. Depending on gene expression, we divided T cells into eight clusters, including cytotoxic T cells (Cytotoxic T), helper T cells (Th), natural killer cell (NK), double negative T cell (DNT), regulatory T cell (Treg), naive T cell (naive T), exhausted T cell (Exhausted T), as well as a special colony of T cells which expressed high level of epithelial-related genes, called epithium_like T cell (epithium_like) (Figure 3A). T

cells from each patient were dispersively distributed (Supplementary Figure 7A). We found large heterogeneity in different samples of T cells. The composition and proportion of T cell showed different characteristics in different disease stages. For example, DNT accounted for larger proportions in NAG and AG stages. It was worth noting that a sample, GN1_1, which belongs to a patient of AG with severe IM, the proportion of T cells was similar to GN1_3, while the latter was similar to AG. To some extent, this proved that AG was intermediate state between NAG and EGC and IM was the state which closed to EGC. For T cell subsets, DNT (CD3+CD4-CD8-T cells) accounted for the highest proportion in the total T cells. Cytotoxic T and exhausted T cells only existed in the samples of EGC and AG, and the proportion of DNT cells gradually decreased during the progression of the disease. Naïve T, Th, NK cells showed an increasing trend (Figure 3B, Supplementary Figure 7B). Meanwhile, we used classical markers to characterize the correctness of the clustering. As can be seen, the marker genes presented specific expression in specific T cells (Figure 3C).

We showed the top 10 DEGs of T cell subsets using heatmap (Figure 3D). The expression of DEGs in heatmap further verified the accuracy of annotation. For cytotoxic T cells, in addition to cytotoxic genes, there were also high expression of KLRC1²⁵, ZNF683²⁶ that interacted with NK cells. There were certain similarities in gene expression of cytotoxic T cells and exhausted T cells. But DUSP4 were higher in exhausted T cells while DUSP2 in cytotoxic T cells. KIR2DL4 which activated cytotoxicity of NK cells were high in exhausted T cells, indicated that both cytotoxic T cells and exhausted T cells could interact with NK cells. DNT concentrated on heat shock protein-related genes, including HSPA1A²⁷, HSPA1B²⁸, DNAJB1²⁹, DNAJA1³⁰. DDX3X is related to the formation of NLRP3 inflammasome³¹. EZR³² and MYH9³³ were in connection with cell polarity such as migration and invasion.

We tracked analysis of total T cells and found no obviously developmental pattern. However, when we just focused on cytotoxic T cells and exhausted T cells, we found the cytotoxic T cells were prior to exhausted T cells and most of the exhausted T cells gathered in the terminal of the branch as a whole, implying the conversion of these

two kinds of cells (Figure 3E). We compared the expression of functional genes between cytotoxic T and exhausted T cells. For cytotoxic genes including IFNG, NKG7, PRF1, GZMA, GZMB, GZMH, there was no difference between the two celltypes. But for immune checkpoint including LAG3, PDCD1, CTLA4 as well as TIGIT, expression of the exhausted T cells was remarkably higher (Figure 3F). The same cell cluster may exhibit diverse states in distinct patients, by finding DEGs, we got the number of cytotoxic T and exhausted T cells DEGs in AG and EGC group.

The total pathway activation patterns were similar between cytotoxic T cells and exhausted T cells, while the latter powered higher activities, except intestinal immune network for IgA production, WNT beta-catenin signal, DNA methylation signaling were mainly enriched in exhausted T cells. Comparing cytotoxic T cells and exhausted T cells in AG groups, the cells in EGC groups presented higher activity in pathways including TRAF6 mediated IRF7 activation pathway, RUNX3 pathway, INF signaling, CXC chemokine receptor signaling and immune response to tumor cells (Supplementary Figure 8B).

Then, we used SPOTlight to locate cytotoxic T cells and exhausted T cells. The distribution of the two types of cells was similar to some extent, but cytotoxic T cells had a wider location (Figure 3G, Supplementary Figure 7F).

When analyzing subsets of T cells, we found that a group of cells that neither CD4 nor CD8 was expressed which we called DNT (Supplementary Figure 8C-E). Meanwhile, the number of DNT was 1890, accounting for 39% of T cells and 8.7% of all cells. Although there were some reports about DNT, the function and effect of DNT were still not clear, and hitherto there was no cover about DNT in GC, then we focused on the DNT in the following.

We explored the difference between the three groups in DNT. The trajectory suggested DNT changed step by step as the disease develops (Supplementary Figure 8F). We checked the gene expression of DNT over pseudotime, and the heat shock protein-related genes such as HSPA9, HSPA1A, HSPA1B, HSP90AB1 as well as DEAD-box helicase genes such as DDX5, DDX3X, DDX3Y were taper off (Supplementary Figure 9A). On the contrary, the phenotype of CD8⁺ T cells

gradually emerged, as the CD8A, GZMK were upregulated. In the later periods of pseudotime, the genes of JCHAIN and IG-related genes highly expressed in plasma cells (Supplementary Figure 9B). The gene expression of DNT pseudotime indicated that transitions to CD8⁺ T cells and plasmas may occur in later development of DNT.

SPOTlight showed that DNT distributed pervasively (Supplementary Figure 8D, Supplementary Figure 9C). But interestingly, areas of lymphocyte aggregation gathered fewer DNT compared to other regions. In the sites of cancer lesions, there existed relatively more DNT (Figure 3G, Supplementary Figure 7F). In combination with the results of monocle, we wildly guessed that DNT may be converted into CD8⁺ T cells to exert the function of inhibiting cancer at the lesion site (Supplementary Figure 8F).

The dysregulation of B cells development during EGC progression

In contrast to T cells, B cells population showed less number and diversity between subgroups (Figure 4A-B). According to the unique marker genes, these cells were divided into B cells (CD19, CD20/MS4A1, FCER2 and MHC II molecules) and plasmas (TNFRSF17, SLAMF7, XBP1, SDC1 and IGHs) (Figure 4E-F). In line with the previous studies that gastric tissues contained massive plasmas, we also found the phenomenon which hitherto with unknown mechanisms³⁴. Spatial location also indicated the diffuse distribution of plasmas while scarce B cells (Figure 4G, Supplementary Figure 11A-D). However, the cell numbers of B cells and plasmas is comparable, respectively, among NAG, AG and EGC groups (Figure 4C-D).

Higher resolution clustering showed there were 7 different subgroups in B cells (Supplementary Figure 10A). B-c0 and B-c5 mainly originated from non-EGC groups while in B-c1, B-c2, B-c3, B-c4 and B-c6, cells from EGC and AG groups predominated (Supplementary Figure 10C). Gene expression analysis showed the similar up-regulated genes between B-c0 and B-c5, DDX5^{35 36}, DDX3X³⁷ and NCL³⁸ were all anti-pathogens and anti-tumor genes. As plenty of oncogenes significantly up-regulated in B-c1 cluster, we inferred it was deteriorating cell populations induced by tumor cells. Then, pathway activity analysis indicating the highly up-regulated

activity of pro-tumor signaling, which also showing the malignant potentials of these cells (Supplementary Figure 10F). For B-c2 cluster, RGS13³⁹, MEF2B⁴⁰, GCSAM⁴¹ and LRMP⁴², essential genes for development of germinal center, were up-regulated. However, FCRL4, usually expressed on the memory B cells under infection, its overexpression also could impede antigen-induced BCR signaling⁴³. Besides, ACTG1⁴⁴ and FAM3C^{45 46}, which were reported to promote tumor development, were also up-regulated in this cluster. For B-c3, it was defined as follicular B cells for the expression of IGHD and FCER2. TCL1A⁴⁷, TNFAIP3⁴⁸, BTG1 and TXNIP⁴⁹, pro-tumor or pro-inflammatory genes, were also expressed in this cluster. For B-c4, it was defined as memory B cells, and the marker gene FABP5 was reported to regulate fatty acid metabolism^{50 51}. In B-c6, SIK2 was reported to regulate metabolism reprogramming through PI3K/AKT/HIF axis, promoting tumor progression; WDR76 could activate fatty metabolism restriction^{52 53}; CNR2 may inhibit antibody response and is reported as central effector of B cells restriction of glucose and energy supply⁵⁴, which in turn indicting that this cluster was lined with the metabolism changes of B cells (Supplementary Figure 10F).

TF analysis identified IRF8, PAX5 and SPIB, the well-known TF essential for development of B cells, CREB3L2 and XBP1 for maturation of plasmas. Moreover, transcriptional activity of EGR3, KLF12, MECP2, HDAC1 and SOX5 were higher in c0-memory B cells and c5-naïve B cells. On the contrary, FOXP1, HMGB1, EGR1 and BCL11A were lower in these clusters but higher in other B cell clusters. It was reported that FOXP1 presented physiological reduction in germinal center B cells, implying that the high activity of FOXP1 leading to infantilization of B cells⁵⁵. HMGB1 could be activated after stimulation of hypoxia or inflammation⁵⁶. BCL11A could exert functions in early stages of B cells^{57 58}. EGR3 could regulate proliferation of B cells through antigen receptor signaling and controlling inflammation⁵⁹. HDAC1 could regulate the G1-to-S phase transition of cell cycle, its ablation blocked B cells at early stages⁶⁰. SOX5 was important for development of memory B cells⁶¹ (Supplementary Figure 10E).

With further dissection of plasmas, we found these cells possessed the evolutionary

trajectory from plasmablasts, shorter-term plasmas to longer-term plasmas (Supplementary Figure 10B, 10D, 10K). Given somatic hypermutation (SHM) and class switch during affinity maturation, plasmas could obtain more variations than B cells, and inferCNV analysis also indicted these differences (Supplementary Figure 10G-J).

Interestingly, spatial feature images showed the diffuse distribution of plasmas and epithelial cells in gastric tissues, on the other hand, the clear clustering was difficult between the two celltypes. Thus, IF was used to figure out the spatial location of them. We found considerable numbers of neighborhood relationships between the two celltypes, especially in EGC patients (Figure 4H). However, there were no illustration for the above phenomenon³⁴, which need further exploration.

As TCR clonotypes were different between NAG, AG and EGC patients, we also identified the BCR information (Supplementary Figure 12A). Diversity of clonotypes were increased from NAG, AG, to EGC groups, (Supplementary Figure 12B-C). Besides, the V-J pairs whatever Heavy chains or Light chains among three groups were unique and different (Supplementary Figure 12D-E).

Characteristics of myeloid cells in the TME of EGC

We then dissected the gene signatures of myeloid subsets revealed in this study (Figure 5A, Supplementary Figure 13A-B). Two mast cell clusters expressed a canonical gene pattern, such as TSPAB1, TPSB2, CPA3, and KIT (Figure 5D, 5G, Supplementary Figure 14C-D), and further distinguished by unique expression of PGA3, NEAT1, HSP90AB1 and TNFAIP3, AREG, MIF, respectively (Figure 5D). Notably, previous studies showed upregulation of TNFAIP3 could promotes gastric cancer development⁶², and AREG could enhance suppressive function of Tregs⁶³. The two clusters showed different enrichment in EGC, AG, and NAG groups (Figure 5B). Mast c0 cluster existed in all three groups more or less evenly, while mast c1 nearly disappeared in NAG groups. The specific distribution may imply different biological roles they functioned. Pathway scores showed mast c0 cluster correlated with KRAS signaling down, intestinal absorption, ABC transporters, and Hedgehog

signaling pathway. Mast c1 cluster tended to become a malignant cell population, and significantly enriched with PI3K-AKT-mTOR, α -TNF, fatty acid metabolism, hypoxia, inflammatory response and IFN- γ signaling (Supplementary Figure 13C).

TF analysis indicated that JUND, MAFF, PRDM1, FOSL2, TCF7L2 and HES1 were activated in macrophages c0 cluster, while the TF activity of FOXO1, FOXO3 and MEF2A were higher in macrophages c1 cluster. As for mast cells, AHR, GATA1, GATA2 were well-known TF, while RELA, EGR3 and ZC3H11A were higher in mast c0 cluster. HIVEP2, ATF4, VDR, ATF5 and IRF4 hold higher activity in Neutrophil/DC cells (Figure 5E).

Two macrophage clusters were identified based on their specific expression of C1QA, C1QB, C1QC, and CD163 (Figure 5D, 5G, Supplementary Figure 14A). Among them, clusters showing marked (macrophage c0 cluster) or slightly (macrophage c1 cluster) skewed distribution in three groups (Figure 5B-C). Expression of MMP9, CXCL2, MMP12 and CTSB were observed in macrophage c0 population, both of which exhibited lower expression in macrophage c1 population (Figure 5D). The remaining clusters were Neutrophil/DC population, characterized by high expression of S100A8, A100A9, G0S2 and LSP1 (Figure 5G, Supplementary Figure 14B). Most of the cluster originated from EGC patients (Figure 5B), and the gene expression as well as activities of enriched pathways were also partial to pro-tumor (Supplementary Figure 13C). Both of macrophages and DC population demonstrated the characteristics of antigen presenting cells (APCs), with abilities of antigen procession and presentation. Apart from this, cell adhesion molecules cams, intestinal immune network for IgA production and complement coagulation cascades signaling were also activated in these cells (Supplementary Figure 13C).

The ligand-receptor interactions of immune cells along with EGC progression

As for the complicated and high-frequency interaction in tissues microenvironment, especially in tumor tissues. We further evaluated the cell interaction among NAG, AG, and EGC groups. The interaction strength and numbers for secreted signaling, cell to cell contact signaling and ECM receptor signaling were all different among three

groups. As contrast to NAG group, the interaction numbers and strengths between AG and EGC groups were comparative, indicating AG patients contained conditions driven early tumor occurrence (Figure 6A, 6C-D, Supplementary Figure 15E-F, G, Supplementary Figure 16A, Supplementary Figure 16A-B). Difference analysis of ligand-receptors indicated that LGALS9_CD44, LGALS9_CD45, LGALS9 - HAVCR2(TIM-3), MIF-CD74/CD44, MIF-CD74/CXCR4 and MDK-NCL were major interactions (Figure 6B, Supplementary Figure 15A-D, Supplementary Figure 16C-E, Supplementary Figure 16C-D). We found LGALS9 was up-regulated in mast cells, macrophages, parietal cells and goblet cells. LGALS9_CD45 reported to inhibit B cell receptor signaling ⁶⁴. LGALS9_CD44 could enhance stability and function of Tregs ⁶⁵. Besides, LGALS9 could interact with HAVCR2 (TIM-3) to protect breast cancer against cytotoxic killing of T cells ⁶⁶, partly by inducing apoptosis or exhaustion of effector T cells ⁶⁷. TIM-3, was reported as a specific surface molecule of Th1 cells and could mediate Th1 immune responses ⁶⁸, its dysregulation could lead to inhibited inflammatory response even tumor evasion ⁶⁹. It could promote CD8⁺ T cell exhaustion and induce production of myeloid-derived suppressor cells (MDSCs) ⁷⁰. In our study, TIM-3 mainly expressed in myeloid cells, in line with the previous reports that TIM-3 expression was predominantly localized to myeloid cells in both human and murine tumors ⁷¹. TIM-3 in mast cells may induce apoptosis ⁷². Macrophage migration inhibitory factor (MIF), an upstream regulator of adaptive and innate immune, but if dysregulated, was a key driver of acute or chronic inflammation and functioned in leukocyte recruitment ⁷³. MIF was occurred in varieties of immune cells among AG and EGC groups. CXCR4, one of receptor of MIF, is highly expressed in EGC group, especially in B cells and T cells population. It was reported that CXCR4 could impeded T-lymphocyte infiltration and immunotherapy in other tumors ⁷⁴. CD74/CD44, as the receptor complex of MIF, its activation could trigger multiple intracellular signal pathway, including ERK pathway, PI3K/AKT signal, NFκB and MAPK pathway ⁷⁵. As the hypoxia pathway was activated in gastritis conditions, it was reported that HIF-1α- and HIF-2α-dependent MIF could regulate function of myeloid cells partly by interacting with CD74/CXCR4 complexes, and

then activate the p38/MAPK and PI3K-AKT signals ⁷⁶. On the other hand, the complex could also function as positive regulators of B cells homeostasis and maturation ⁷⁷. Midkine (MDK) was reported to support progression of gastric cancer ⁷⁸, partly by inhibiting expression of Granzyme B (GZMB) and cytotoxicity of NK cells ⁷⁹, as well as increased epidermal growth factor receptor (EGFR) signaling under hypoxia through interaction with surface nucleolin (NCL) ⁸⁰. It was also identified to contribute to multidrug resistance in GC cells ⁸¹. Besides, MDK rewires immunosuppressive environment in melanoma and gallbladder cancer ^{82 83}. In this study, we found that MDK is overexpressed in EGC and AG patients comparing with NAG ones. One of receptors of MDK was NCL, a major nucleolar protein of growing cells, also functioned as a cell surface receptor to shuttle between cytoplasm and nucleus, thus providing a mechanism for extracellular regulation of nuclear events ⁸⁴.

Discussion

In order to explore the immune microenvironment changes from gastritis to early gastric cancer, we firstly dissected the immune cells component and spatial structures by comprehensively use of single-cell sequencing and spatial transcriptome technologies. Although there were several gastric cancer-related studies previously, they either focused on the malignant cells, or other un-immune cells in gastric cancer just at single cell resolution ^{7-11 85}, or with other spatial methods such as DSP instead of ST³⁴.

Global pathway enrichment and spatial regions analysis indicated that several pathways such as hypoxia conditions, reactive oxygen species pathway and fatty metabolism were activated in AG stages, all benefited to the progression of EGC and form of inhibitory immune microenvironment, comparing with the NAG group. While MYC, INF α , IL-1 and TNF signaling, all tumor-related pathways were activated in EGC groups. The overall changes implied the heterogeneity of the compositions. Then in this study, multiple immune cells, including T cells, B cells and myeloid cells were found and they possessed different states and expression patterns among EGC,

AG and NAG patients.

When the gastric disease progressed from NAG to EGC, the balance between effector and inhibitory states of T cells were impaired gradually. Specifically, exhausted T cells and cytotoxic T cells were emerged in AG and EGC stages, which were non-existent in NAG state, and immune responses activated obviously. This was consistent with the previous work that early tumor progression was accompanied with immune response activation. In addition to the existence of exhausted T cells, another natural inhibitory role in TME, Tregs, up-regulated transcriptional activity of HIVEP1 and STAT3 in NAG stage, while enhanced activities of IRF4 and IKZF2 in AG state, and possessed higher transcriptional activities of FOXP3, BATF, and PRDM1 in EGC period. Among them, HIVEP1 was a newly found TF in Tregs, it was just reported highly activated in TNFRSF9+ Treg cells, at a late part of the trajectory ⁸⁶, presenting a different pattern from our results. As previous reports, PRDM1 usually up-regulated transcriptional activity in advanced or end-stage, implying worse immune effects ^{87 88}. These changes of transcriptional factors activity during progression from NAG to EGC may indicate that the inhibitory effects of T cells were one of the contributing factors of early gastric cancer. The above also showed that the exhausted or inhibitory immune responses have begun from AG patients, implied the preventive measures for EGC maybe should enforce from this stage. Besides, we found a rare celltype in T cells, the DNT cells, which occupied a large portion but has uncertain marker genes to defined them, which was in accordance with the previous reports ⁸⁹⁻⁹¹. Time trajectory showed the continuous evolution of DNT from NAG, AG, to EGC, without enough experimental verification, thus, the significances of their existence were unclear, which was worth further investigation in the future.

Then, the same as T cells, dysregulation also emerged in B cells. The higher transcriptional activity of FOXP1 and BCL11A in clusters (B-c1, B-c2, B-c3, B-c4 and B-c6) that enriched more in AG and EGC groups, represented the infantilization of B cells ^{55 57 58}. Besides, the transcriptional inhibition of SOX5 indicated the loss of mature memory B cells ⁶¹, as well as maturation disorders for the lower activities of EGR3 ⁵⁹ and HDAC1⁶⁰, which positively regulate cell cycles. Comparing B cells,

plasmas were rich in gastric tissues, the same as previous reports³⁴. We have found an interesting phenomenon that plasmas were co-localized or accompanied with epithelial cells, especially in EGC groups, which confirmed by ST images of marker genes expression and IF staining. However, the cellular interaction could not find out hints for the phenomenon, showing possibilities for further exploration in depth.

There were two malignant clusters, Mast c1 and macrophage c0, mainly enriched in AG and EGC groups, which expressed high level genes that promoted tumor progression and inhibited effective immune responses. As for AREG up-regulated in Mast c1, could interact with EGFR to enforce the function of Tregs^{63 92}. LGALS9, mainly expressed in myeloid cells, could interact with TIM-3 to exert immune suppressive roles^{66 69}. Besides, both MIF-CD74/CXCR4 and MDK-NCL could activate in hypoxia condition, and regulate functions of T cells or myeloid cells, respectively^{76 80}.

The strength of this study came from application of combined technologies: spatial transcriptomics, scRNA-seq and VDJ analysis. Visium ST has allowed us to determine the location of cell types within different tissues; although scRNA-seq could not acquire enough depths of gene expression just like the common RNA-Seq, it has allowed us to confirm the known celltypes in immune microenvironment during progression of gastric cancer, as well as separate rare cell types such as Schwann cells and DNT cells.

However, there were also several limitations in this study: (1) Given the samples incorporated in this study were not enough, more samples were needed. (2) We mainly focused on the immune cells, other un-immune cells which were also extremely important for EGC progression, reported previously⁷⁻¹¹, were not clarified in detail, it will be presented in latter works. (3) The presented conclusions mostly depended on the computationally analysis, and more functional experiment were needed to convey for further identification. (4) The exploration of molecular mechanisms was not enough, which needed more experimental outcomes as bridges to every hypothesis.

Overall, we characterized the dysregulation of immune microenvironment

changes both in single-cells and spatial transcriptome resolution, the multi-omics data presented in this study offer a comprehensive view of immune cell types, states change and locations within the gastric tissues during the progression from NAG, AG to EGC, advancing our understanding of the composition and immunity of different gastric states, offering diagnostic and preventive thoughts for EGC.

Competing interests: The authors declare that they have no conflict of interests.

Acknowledgement

We thank Department of pathology in the Second Hospital of Shandong University for species and technical assistance.

Reference

- 1 Incidence of gastric cancer among patients with gastric precancerous lesions: observational cohort study in a low risk Western population. *BMJ (Clinical research ed)* 2015;**351**:h4134.
- 2 Oka S, Tanaka S, Kaneko I, *et al.* Advantage of endoscopic submucosal dissection compared with EMR for early gastric cancer. *Gastrointestinal endoscopy* 2006;**64**:877-83.
- 3 Suzuki H, Oda I, Abe S, *et al.* High rate of 5-year survival among patients with early gastric cancer undergoing curative endoscopic submucosal dissection. *Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association* 2016;**19**:198-205.
- 4 Smyth EC, Nilsson M, Grabsch HI, *et al.* Gastric cancer. *Lancet (London, England)* 2020;**396**:635-48.
- 5 Takizawa K, Ono H, Hasuike N, *et al.* A nonrandomized, single-arm confirmatory trial of expanded endoscopic submucosal dissection indication for undifferentiated early gastric cancer: Japan Clinical Oncology Group study (JCOG1009/1010). *Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association* 2021;**24**:479-91.
- 6 Hinshaw DC, Shevde LA. The Tumor Microenvironment Innately Modulates Cancer Progression. *Cancer research* 2019;**79**:4557-66.
- 7 Zhang M, Hu S, Min M, *et al.* Dissecting transcriptional heterogeneity in primary gastric adenocarcinoma by single cell RNA sequencing. *Gut* 2021;**70**:464-75.
- 8 Zhang P, Yang M, Zhang Y, *et al.* Dissecting the Single-Cell Transcriptome Network Underlying Gastric Premalignant Lesions and Early Gastric Cancer. *Cell reports* 2019;**27**:1934-47.e5.
- 9 Han S, Fink J, Jörg DJ, *et al.* Defining the Identity and Dynamics of Adult Gastric Isthmus Stem Cells. *Cell stem cell* 2019;**25**:342-56.e7.

- 10 Bockerstett KA, Lewis SA, Noto CN, *et al.* Single-Cell Transcriptional Analyses Identify Lineage-Specific Epithelial Responses to Inflammation and Metaplastic Development in the Gastric Corpus. *Gastroenterology* 2020;**159**:2116-29.e4.
- 11 Bockerstett KA, Lewis SA, Wolf KJ, *et al.* Single-cell transcriptional analyses of spasmolytic polypeptide-expressing metaplasia arising from acute drug injury and chronic inflammation in the stomach. *Gut* 2020;**69**:1027-38.
- 12 Wang A, Nie S, Lv Z, *et al.* Infiltration of Immunoinflammatory Cells and Related Chemokine/Interleukin Expression in Different Gastric Immune Microenvironments. *Journal of immunology research* 2020;**2020**:2450569.
- 13 De Salvo C, Pastorelli L, Petersen CP, *et al.* Interleukin 33 Triggers Early Eosinophil-Dependent Events Leading to Metaplasia in a Chronic Model of Gastritis-Prone Mice. *Gastroenterology* 2021;**160**:302-16.e7.
- 14 Sathe A, Grimes SM, Lau BT, *et al.* Single-Cell Genomic Characterization Reveals the Cellular Reprogramming of the Gastric Tumor Microenvironment. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2020;**26**:2640-53.
- 15 Yin H, Guo R, Zhang H, *et al.* A Dynamic Transcriptome Map of Different Tissue Microenvironment Cells Identified During Gastric Cancer Development Using Single-Cell RNA Sequencing. *Frontiers in immunology* 2021;**12**:728169.
- 16 Wang R, Dang M, Harada K, *et al.* Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma. *Nature medicine* 2021;**27**:141-51.
- 17 Longo SK, Guo MG, Ji AL, *et al.* Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nature reviews Genetics* 2021;**22**:627-44.
- 18 Moncada R, Barkley D, Wagner F, *et al.* Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nature biotechnology* 2020;**38**:333-42.

- 19 Ji AL, Rubin AJ, Thrane K, *et al.* Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma. *Cell* 2020;**182**:497-514.e22.
- 20 Wu Y, Yang S, Ma J, *et al.* Spatiotemporal Immune Landscape of Colorectal Cancer Liver Metastasis at Single-Cell Level. *Cancer discovery* 2022;**12**:134-53.
- 21 Chan JM, Quintanal-Villalonga Á, Gao VR, *et al.* Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. *Cancer cell* 2021;**39**:1479-96.e18.
- 22 Jin S, Guerrero-Juarez CF, Zhang L, *et al.* Inference and analysis of cell-cell communication using CellChat. *Nature communications* 2021;**12**:1088.
- 23 Patel AP, Tirosh I, Trombetta JJ, *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science (New York, NY)* 2014;**344**:1396-401.
- 24 Elosua-Bayes M, Nieto P, Mereu E, *et al.* SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic acids research* 2021;**49**:e50.
- 25 Chen Y, Xin Z, Huang L, *et al.* CD8(+) T Cells Form the Predominant Subset of NKG2A(+) Cells in Human Lung Cancer. *Frontiers in immunology* 2019;**10**:3002.
- 26 Post M, Cuapio A, Osl M, *et al.* The Transcription Factor ZNF683/HOBIT Regulates Human NK-Cell Development. *Frontiers in immunology* 2017;**8**:535.
- 27 Wu S, Pei Q, Ni W, *et al.* HSPA1A Protects Cells from Thermal Stress by Impeding ESCRT-0-Mediated Autophagic Flux in Epidermal Thermoresistance. *The Journal of investigative dermatology* 2021;**141**:48-58.e3.
- 28 Li Y, Yu S, Gu G, *et al.* Polymorphisms of heat shock protein 70 genes (HSPA1A, HSPA1B and HSPA1L) and susceptibility of noise-induced hearing loss in a Chinese population: A case-control study. *PloS one*

- 2017;**12**:e0171722.
- 29 Cui X, Choi HK, Choi YS, *et al.* DNAJB1 destabilizes PDCD5 to suppress p53-mediated apoptosis. *Cancer letters* 2015;**357**:307-15.
- 30 Kaida A, Yamamoto S, Parrales A, *et al.* DNAJA1 promotes cancer metastasis through interaction with mutant p53. *Oncogene* 2021;**40**:5013-25.
- 31 Mo J, Liang H, Su C, *et al.* DDX3X: structure, physiologic functions and cancer. *Molecular cancer* 2021;**20**:38.
- 32 Xu J, Zhang W. EZR promotes pancreatic cancer proliferation and metastasis by activating FAK/AKT signaling pathway. *Cancer cell international* 2021;**21**:521.
- 33 Li Y, Liu X, Lin X, *et al.* Chemical compound cinobufotalin potently induces FOXO1-stimulated cisplatin sensitivity by antagonizing its binding partner MYH9. *Signal transduction and targeted therapy* 2019;**4**:48.
- 34 Kumar V, Ramnarayanan K, Sundar R, *et al.* Single-Cell Atlas of Lineage States, Tumor Microenvironment, and Subtype-Specific Expression Programs in Gastric Cancer. *Cancer discovery* 2021.
- 35 Sun J, Wu G, Pastor F, *et al.* RNA helicase DDX5 enables STAT1 mRNA translation and interferon signalling in hepatitis B virus replicating hepatocytes. *Gut* 2021.
- 36 Dixon CR, Malik P, de Las Heras JI, *et al.* STING nuclear partners contribute to innate immune signaling responses. *iScience* 2021;**24**:103055.
- 37 Soulat D, Bürckstümmer T, Westermayer S, *et al.* The DEAD-box helicase DDX3X is a critical component of the TANK-binding kinase 1-dependent innate immune response. *The EMBO journal* 2008;**27**:2135-46.
- 38 Shi H, Huang Y, Zhou H, *et al.* Nucleolin is a receptor that mediates antiangiogenic and antitumor activity of endostatin. *Blood* 2007;**110**:2899-906.
- 39 Shi GX, Harrison K, Wilson GL, *et al.* RGS13 regulates germinal center B lymphocytes responsiveness to CXC chemokine ligand (CXCL)12 and CXCL13. *Journal of immunology (Baltimore, Md : 1950)* 2002;**169**:2507-15.

- 40 Chu CS, Hellmuth JC, Singh R, *et al.* Unique Immune Cell Coactivators Specify Locus Control Region Function and Cell Stage. *Molecular cell* 2020;**80**:845-61.e10.
- 41 Raboso-Gallego J, Casado-García A, Jiang X, *et al.* Conditional expression of HGAL leads to the development of diffuse large B-cell lymphoma in mice. *Blood* 2021;**137**:1741-53.
- 42 Tedoldi S, Paterson JC, Cordell J, *et al.* Jaw1/LRMP, a germinal centre-associated marker for the immunohistological study of B-cell lymphomas. *The Journal of pathology* 2006;**209**:454-63.
- 43 Sohn HW, Krueger PD, Davis RS, *et al.* FcRL4 acts as an adaptive to innate molecular switch dampening BCR signaling and enhancing TLR signaling. *Blood* 2011;**118**:6332-41.
- 44 Richter C, Mayhew D, Rennhack JP, *et al.* Genomic Amplification and Functional Dependency of the Gamma Actin Gene ACTG1 in Uterine Cancer. *International journal of molecular sciences* 2020;**21**.
- 45 Schmidt U, Heller G, Timelthaler G, *et al.* The FAM3C locus that encodes interleukin-like EMT inducer (ILEI) is frequently co-amplified in MET-amplified cancers and contributes to invasiveness. *Journal of experimental & clinical cancer research : CR* 2021;**40**:69.
- 46 Shi M, Duan G, Nie S, *et al.* Elevated FAM3C promotes cell epithelial-mesenchymal transition and cell migration in gastric cancer. *OncoTargets and therapy* 2018;**11**:8491-505.
- 47 Brinas F, Danger R, Brouard S. TCL1A, B Cell Regulation and Tolerance in Renal Transplantation. *Cells* 2021;**10**.
- 48 Zhang J, Grubor V, Love CL, *et al.* Genetic heterogeneity of diffuse large B-cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America* 2013;**110**:1398-403.
- 49 Wang CY, Xu Y, Wang X, *et al.* Dl-3-n-Butylphthalide Inhibits NLRP3 Inflammasome and Mitigates Alzheimer's-Like Pathology via Nrf2-TXNIP-TrX Axis. *Antioxidants & redox signaling* 2019;**30**:1411-31.

- 50 Lin R, Zhang H, Yuan Y, *et al.* Fatty Acid Oxidation Controls CD8(+) Tissue-Resident Memory T-cell Survival in Gastric Adenocarcinoma. *Cancer immunology research* 2020;**8**:479-92.
- 51 Pan Y, Tian T, Park CO, *et al.* Survival of tissue-resident memory T cells requires exogenous lipid uptake and metabolism. *Nature* 2017;**543**:252-6.
- 52 Jiang Y, Mao C, Yang R, *et al.* EGLN1/c-Myc Induced Lymphoid-Specific Helicase Inhibits Ferroptosis through Lipid Metabolic Gene Expression Changes. *Theranostics* 2017;**7**:3293-305.
- 53 Huang D, Li Q, Sun X, *et al.* CRL4(DCAF8) dependent opposing stability control over the chromatin remodeler LSH orchestrates epigenetic dynamics in ferroptosis. *Cell death and differentiation* 2021;**28**:1593-609.
- 54 Chan LN, Chen Z, Braas D, *et al.* Metabolic gatekeeper function of B-lymphoid transcription factors. *Nature* 2017;**542**:479-83.
- 55 Sagardoy A, Martinez-Ferrandis JI, Roa S, *et al.* Downregulation of FOXP1 is required during germinal center B-cell function. *Blood* 2013;**121**:4311-20.
- 56 Zhang Y, Pan N, Sheng Y, *et al.* Hypoxia enhances IL-10-producing B cell generation through upregulating high-mobility group B1 on tumor cell-released autophagosomes. *Immunology letters* 2019;**216**:36-42.
- 57 Yu Y, Wang J, Khaled W, *et al.* Bcl11a is essential for lymphoid development and negatively regulates p53. *The Journal of experimental medicine* 2012;**209**:2467-83.
- 58 Sunami Y, Yokoyama T, Yoshino S, *et al.* BCL11A promotes myeloid leukemogenesis by repressing PU.1 target genes. *Blood advances* 2021.
- 59 Li S, Miao T, Sebastian M, *et al.* The transcription factors Egr2 and Egr3 are essential for the control of inflammation and antigen-induced proliferation of B and T cells. *Immunity* 2012;**37**:685-96.
- 60 Yamaguchi T, Cubizolles F, Zhang Y, *et al.* Histone deacetylases 1 and 2 act in concert to promote the G1-to-S progression. *Genes & development* 2010;**24**:455-69.
- 61 Ehrhardt GR, Hijikata A, Kitamura H, *et al.* Discriminating gene expression

- profiles of memory B cell subpopulations. *The Journal of experimental medicine* 2008;**205**:1807-17.
- 62 Du B, Liu M, Li C, *et al.* The potential role of TNFAIP3 in malignant transformation of gastric carcinoma. *Pathology, research and practice* 2019;**215**:152471.
- 63 Zaiss DM, van Loosdregt J, Gorlani A, *et al.* Amphiregulin enhances regulatory T cell-suppressive function via the epidermal growth factor receptor. *Immunity* 2013;**38**:275-84.
- 64 Giovannone N, Liang J, Antonopoulos A, *et al.* Galectin-9 suppresses B cell receptor signaling and is regulated by I-branching of N-glycans. *Nature communications* 2018;**9**:3287.
- 65 Wu C, Thalhamer T, Franca RF, *et al.* Galectin-9-CD44 interaction enhances stability and function of adaptive regulatory T cells. *Immunity* 2014;**41**:270-82.
- 66 Yasinska IM, Sakhnevych SS, Pavlova L, *et al.* The Tim-3-Galectin-9 Pathway and Its Regulatory Mechanisms in Human Breast Cancer. *Frontiers in immunology* 2019;**10**:1594.
- 67 Hu XH, Tang MX, Mor G, *et al.* Tim-3: Expression on immune cells and roles at the maternal-fetal interface. *Journal of reproductive immunology* 2016;**118**:92-9.
- 68 Zhu C, Anderson AC, Schubart A, *et al.* The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. *Nature immunology* 2005;**6**:1245-52.
- 69 Han G, Chen G, Shen B, *et al.* Tim-3: an activation marker and activation limiter of innate immune cells. *Frontiers in immunology* 2013;**4**:449.
- 70 Fujihara S, Mori H, Kobara H, *et al.* Galectin-9 in cancer therapy. *Recent patents on endocrine, metabolic & immune drug discovery* 2013;**7**:130-7.
- 71 de Mingo Pulido Á, Gardner A, Hiebler S, *et al.* TIM-3 Regulates CD103(+) Dendritic Cell Function and Response to Chemotherapy in Breast Cancer. *Cancer cell* 2018;**33**:60-74.e6.

- 72 Nakae S, Iikura M, Suto H, *et al.* TIM-1 and TIM-3 enhancement of Th2 cytokine production by mast cells. *Blood* 2007;**110**:2565-8.
- 73 Bernhagen J, Krohn R, Lue H, *et al.* MIF is a noncognate ligand of CXC chemokine receptors in inflammatory and atherogenic cell recruitment. *Nature medicine* 2007;**13**:587-96.
- 74 Chen IX, Chauhan VP, Posada J, *et al.* Blocking CXCR4 alleviates desmoplasia, increases T-lymphocyte infiltration, and improves immunotherapy in metastatic breast cancer. *Proceedings of the National Academy of Sciences of the United States of America* 2019;**116**:4558-66.
- 75 Su H, Na N, Zhang X, *et al.* The biological function and significance of CD74 in immune diseases. *Inflammation research : official journal of the European Histamine Research Society [et al]* 2017;**66**:209-16.
- 76 Zhu G, Tang Y, Geng N, *et al.* HIF- α /MIF and NF- κ B/IL-6 axes contribute to the recruitment of CD11b+Gr-1+ myeloid cells in hypoxic microenvironment of HNSCC. *Neoplasia (New York, NY)* 2014;**16**:168-79.
- 77 Gore Y, Starlets D, Maharshak N, *et al.* Macrophage migration inhibitory factor induces B cell survival by activation of a CD74-CD44 receptor complex. *The Journal of biological chemistry* 2008;**283**:2784-92.
- 78 Xu Y, Qu X, Zhang X, *et al.* Midkine positively regulates the proliferation of human gastric cancer cells. *Cancer letters* 2009;**279**:137-44.
- 79 Zhao S, Wang H, Nie Y, *et al.* Midkine upregulates MICA/B expression in human gastric cancer cells and decreases natural killer cell cytotoxicity. *Cancer immunology, immunotherapy : CII* 2012;**61**:1745-53.
- 80 Kinoshita D, Shishido T, Takahashi T, *et al.* Growth Factor Midkine Aggravates Pulmonary Arterial Hypertension via Surface Nucleolin. *Scientific reports* 2020;**10**:10345.
- 81 Kang HC, Kim IJ, Park JH, *et al.* Identification of genes with differential expression in acquired drug-resistant gastric cancer cells using high-density oligonucleotide microarrays. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2004;**10**:272-84.

- 82 Cerezo-Wallis D, Contreras-Alcalde M, Troulé K, *et al.* Midkine rewires the melanoma microenvironment toward a tolerogenic and immune-resistant state. *Nature medicine* 2020;**26**:1865-77.
- 83 Zhang Y, Zuo C, Liu L, *et al.* Single-cell RNA-sequencing atlas reveals an MDK-dependent immunosuppressive environment in ErbB pathway-mutated gallbladder cancer. *Journal of hepatology* 2021;**75**:1128-41.
- 84 Srivastava M, Pollard HB. Molecular dissection of nucleolin's role in growth and cell proliferation: new insights. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 1999;**13**:1911-22.
- 85 Fu K, Hui B, Wang Q, *et al.* Single-cell RNA sequencing of immune cells in gastric cancer patients. *Aging* 2020;**12**:2747-63.
- 86 Zheng L, Qin S, Si W, *et al.* Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science (New York, NY)* 2021;**374**:abe6474.
- 87 Cretney E, Leung PS, Trezise S, *et al.* Characterization of Blimp-1 function in effector regulatory T cells. *Journal of autoimmunity* 2018;**91**:73-82.
- 88 Cretney E, Xin A, Shi W, *et al.* The transcription factors Blimp-1 and IRF4 jointly control the differentiation and function of effector regulatory T cells. *Nature immunology* 2011;**12**:304-11.
- 89 Sun G, Zhao X, Li M, *et al.* CD4 derived double negative T cells prevent the development and progression of nonalcoholic steatohepatitis. *Nature communications* 2021;**12**:650.
- 90 Zhang ZX, Young K, Zhang L. CD3+CD4-CD8- alphabeta-TCR+ T cell as immune regulatory cell. *Journal of molecular medicine (Berlin, Germany)* 2001;**79**:419-27.
- 91 Fischer K, Voelkl S, Heymann J, *et al.* Isolation and characterization of human antigen-specific TCR alpha beta+ CD4(-)CD8- double-negative regulatory T cells. *Blood* 2005;**105**:2828-35.
- 92 Meulenbroeks C, van Weelden H, Schwartz C, *et al.* Basophil-derived amphiregulin is essential for UVB irradiation-induced immune suppression. *The Journal of investigative dermatology* 2015;**135**:222-8.

Figure 1 The abstract of spatial transcriptomes sequencing.

(A) The flowcharts of scRNA-seq, ST and VDJ sequencing. (B) The TSNE plot for all clusters of ST. (C) The TSNE plot for NAG, AG and EGC three groups of ST. (D) The time trajectory analysis from NAG, AG to EGC of ST. (E) The time trajectory changes of DEGs from NAG to EGC of ST. (F) The images with circled regions of ST sequencing.

Figure 2 The abstract of single-cell sequencing.

(A) The TSNE plot for all identified celltypes of scRNA-seq. (B) The TSNE plot for 6 samples of scRNA-seq. (C) The TSNE plot for NAG, AG and EGC three groups of scRNA-seq. (D) The marker genes for all identified celltypes of scRNA-seq. (E) The cell proportions (left) among 6 samples and 3 groups, as well as cell numbers (right) of all identified celltypes of scRNA-seq. (F) The spatial locations of identified celltypes (epithelium cells, B cells, plasmas, T cells and smooth cells) in 3 ST samples (C8, N21-1 and N22-1).

Figure 3 The characteristics of subclusters in T cells.

(A) The TSNE plot for subclusters of T cells. (B) The cell proportions (left) among 6 samples and 3 groups, as well as cell numbers (right) of subclusters of T cells. (C) The TSNE feature plots for marker genes of subclusters in T cells. (D) The top 10 marker genes of each subcluster of T cells. (E) The time trajectory between Cytotoxic and exhausted T cells. (F) The relative expression of markers genes in Cytotoxic and exhausted T cells. (G) The spatial locations of Cytotoxic and exhausted T cells in AG (N21.1 and N8.1) and EGC (C7 and C8) groups.

Figure 4 The dysregulation of B cell subclusters.

(A) The TSNE plot for B cells and plasmas. (B) The TSNE plot for B cells and plasmas among 3 groups. (C) The cell proportions among 3 groups of B cells and plasmas. (D) The cell proportions among 6 samples of B cells and plasmas. (E) The marker genes of B cells and plasmas. (F) The TSNE feature plots of marker genes of B cells and plasmas. (G) The spatial location of marker genes of B cells and plasmas. (H) The IF staining of plasmas and epithelium cells (Red: CD138, represented plasmas; Green: PanCK represented epithelial cells; Blue: DAPI, represented cell nucleus).

Figure 5 The features of myeloid cells.

(A) The TSNE plot for myeloid cells. (B) The cell proportions among 3 groups of myeloid cells. (C) The cell proportions among 6 samples of myeloid cells. (D) The marker genes of myeloid cells. (E) The transcriptional factors of myeloid cells. (F) The TSNE feature plots of marker genes of myeloid cells. (G) The spatial location of marker genes of myeloid cells.

Figure 6 The cellular interaction among identified celltypes.

(A) The cellular interaction numbers and strengths among 3 groups. (B) The ligand-receptor pairs from epithelium cell to other celltypes. (C) The cellular interaction numbers of secreted signaling. (D) The cellular interaction strengths of secreted signaling. (E) The interaction involved pathways between AG and EGC (left), NAG and EGC (median), as well as NAG and AG (right) of secreted signaling.

Supplementary Figure 1 The pathway activities of samples used for ST, related to Figure 1.

(A) The cell proportions among 3 groups of ST data. (B) The cell proportions among 12 samples of ST data. (C) The TSNE plot for ST data among 12 samples. (D) The key pathways enriched in each sample used for ST. (E) The key pathways enriched in each group of ST. The pathway activities of (F) KRAS down signaling down pathway and (G) Inflammatory response pathway represented by spots among 12 sample used for ST.

Supplementary Figure 2 The activities changes of 4 key pathway in 12 samples used for ST, related to Figure 1.

The pathway activities of (A) Hypoxia, (B) MYC signaling, (C) DNA repair, and (D) Interferon alpha response pathway represented by spots among 12 sample used for ST.

Supplementary Figure 3 The identification of celltypes by MIA methods, related to Figure 1.

(A) The identification of celltypes by MIA methods.

Supplementary Figure 4 The evaluation of CNV in gastric TME, related to Figure 1.

(A) The TSNE plot of all clusters of ST data split by 3 groups. (B) The evaluation of CNV changes among 3 groups by inferCNV. The NAG group was used as control. (C) The CNV score in each cluster of ST data. (D) The CNV score in each sample of ST data. (E) The CNV score in each group of ST data.

Supplementary Figure 5 The similarity of specific regions in each ST image, related to Figure 1.

(A) The similarity of specific regions in each ST image.

Supplementary Figure 6 The abstract of single-cell sequencing, related to Figure 2.

(A) The pathological information of patients used in this study. (B) The TSNE feature plots of marker genes for identified celltypes. (C) The cell proportions of identified celltypes among 6 patients used for scRNA-seq. (D) The expression of TFF2 among NAG, AG and EGC groups. (E) Average proportion of each spot in identified celltypes. (F) The celltype location in spatial images (C8, N21-1 and N22-1) of fibroblasts, parietal cells, endothelium, mast cells, endocrine cells and pericytes.

Supplementary Figure 7 The features of T cells, related to Figure 3.

(A) The pathological information of patients used in this study. (B) The TSNE feature plots of marker genes for identified celltypes. (C) The cell proportions of identified celltypes among 6 patients used for scRNA-seq. (D) The expression of TFF2 among NAG, AG and EGC groups. (E) Average proportion of each spot in identified celltypes. (F) The celltype location in spatial images (C8, N21-1 and N22-1) of fibroblasts, parietal cells, endothelium, mast cells, endocrine cells and pericytes.

Supplementary Figure 8 The features of specific subclusters of T cell and identification of DNT cells, related to Figure 3.

(A) The expression of selected genes showed in different clusters of Tregs, cytotoxic and exhausted T cells grouped by NAG, AG and EGC. (B) The pathways enriched in different clusters of Tregs, cytotoxic and exhausted T cells grouped by NAG, AG and EGC. (C) The highlighted DNT cells among all T cells. (D) Spatial location of DNT cells in N22.0, N21.1 and C7. (E) The definition of DNT cells with rare expression of CD4 or CD8 markers. (F) The different states used by time trajectory analysis for DNT cells.

Supplementary Figure 9 The DEGs of DNT cells, related to Figure 3.

(A) The DEGs followed the changes of time trajectory. (B) The gene expression along with the changes of time trajectory. (C) Spatial location of DNT cells in NAG (N28.0, N26.0 and N22.1), AG (N8.1, N24.1 and N34.1) and EGC (C8, C3 and C9) groups.

Supplementary Figure 10 The pathway and transcriptional activities of subclusters in B cells, related to Figure 4.

(A) The TSNE plot of subclusters in B cells. (B) The TSNE plot of subclusters in plasmas. (C) The proportion of subclusters in B cells. (D) The proportion of subclusters in plasmas. (E) The transcriptional factors in subclusters of B cells and plasmas. (F) The pathway activities in subclusters of B cells and plasmas. (G) The CNV of B cells estimated by inferCNV. (H) The CNV scores among groups, (I) samples and between (J) B cells and plasmas. (K) The time trajectory of subclusters in plasmas.

Supplementary Figure 11 Spatial location and expression for marker genes of B cells, related to Figure 4.

(A) The spatial location and expression for MS4A1 (CD20), (B) CD19, (C) SDC1, and (D) TNFRSF17 among all 12 samples.

Supplementary Figure 12 The BCR changes of B cells with the progression of EGC, related to Figure 4.

(A) The cells with BCR were highlighted on all cells and B cells. (B) The BCR diversity among NAG, AG and EGC groups. (C) The BCR diversity among 6 samples. (D) The V-J pairs of BCR heavy chains (H represented the heavy chains) and (E) light chains (LK represented the two light chains)

Supplementary Figure 13 The pathway activities of myeloid cells, related to Figure 5.

(A) The TSNE plot of myeloid cells among 3 groups. (B) The TSNE plot of myeloid cells among 6 samples. (C) The pathway activities of different subclusters of myeloid cells.

Supplementary Figure 14 Spatial location and expression for marker genes of myeloid cells, related to Figure 5.

(A) The spatial location and expression for C1QA, (B) S100A8, (C) TPSAB1, and (D) CPA3 among all 12 samples.

Supplementary Figure 15 The cellular interaction of secreted signaling and cell-cell contact signaling among specific cells, related to Figure 6.

(A) The cellular interaction between T cells and other cells about secreted signaling. (B) The cellular interaction between B cells and other cells about secreted signaling. (C) The cellular interaction between fibroblasts and other cells about secreted signaling. (D) The cellular interaction between parietal cells and other cells about secreted signaling. (E) The interaction numbers and strengths of cell-cell contact signaling and (F) ECM signaling. (G) The cellular interaction strengths of cell-cell contact signaling for each pair among NAG, AG and EGC groups.

Supplementary Figure 16 The cellular interaction of cell-cell contact signaling among specific cells, related to Figure 6.

(A) The cellular interaction numbers of cell-cell contact signaling for each pair among NAG, AG and EGC groups. (B) The interaction involved pathways between AG and EGC (left), NAG and EGC (median), as well as NAG and AG (right) of cell-cell contact signaling. (C) The cellular interaction between T cells and other cells about cell-cell contact signaling. (D) The cellular interaction between chief cells and other cells about cell-cell contact signaling. (E) The cellular interaction between endocrine cells and other cells about cell-cell contact signaling.

Supplementary Figure 17 The cellular interaction of ECM signaling among specific cells, related to Figure 6.

(A) The cellular interaction strengths of ECM signaling for each pair among NAG, AG and EGC groups. (B) The cellular interaction numbers of ECM signaling for each pair among NAG, AG and EGC groups. (C) The cellular interaction between fibroblasts and other cells about cell-cell contact signaling. (D) The cellular interaction between pericytes and other cells about cell-cell contact signaling.