注：Method 描述模板仅供参考，请根据实际情况修改使用。

Single-cell RNA-seq data preprocessing:

The Cell Ranger software pipeline (version 5.0.0) provided by 10×Genomics was used to demultiplex cellular barcodes, map reads to the genome and transcriptome using the STAR aligner, and down-sample reads as required to generate normalized aggregate data across samples, producing a matrix of gene counts versus cells. We processed the unique molecular identifier (UMI) count matrix using the R package Seurat[1] (version 3.1.1). To remove low quality cells and likely multiplet captures, which is a major concern in microdroplet-based experiments, we applied a criteria to filter out cells with UMI/gene numbers out of the limit of mean value +/- 2 fold of standard deviations assuming a Guassian distribution of each cells' UMI/gene numbers. Following visual inspection of the distribution of cells by the fraction of mitochondrial genes expressed, we further discarded low-quality cells where >10% of the counts belonged to mitochondrial genes. Additionally, we applied DoubletFinder package[2] (version 2.0.2) to identify potential doublet. After applying these QC criteria, xxxx single cells were included in downstream analyses. Library size normalization was performed with NormalizeData function in Seurat[1] to obtain the normalized count. Specifically, the global-scaling normalization method "LogNormalize" normalized the gene expression measurements for each cell by the total expression, multiplied by a scaling factor (10,000 by default), and the results were logtransformed.

Top variable genes across single cells were identified using the method described in Macosko et al[3]. The most variable genes were selected using FindVariableGenes function(mean.function = FastExpMean, dispersion.function = FastLogVMR) in Seurat[1]. Principal component analysis (PCA) was performed to reduce the dimensionality with RunPCA function in Seurat[1]. Graph-based clustering was performed to cluster cells according to their gene expression profile using the FindClusters function in Seurat[1]. Cells were visualized using a 2-dimensional t-distributed stochastic neighbor embedding (t-SNE) algorithm with the RunTSNE function in Seurat[1]. We used the FindAllMarkers function（test.use = bimod） in Seurat[1] to identify marker genes of each cluster. For a given cluster, FindAllMarkers

批注 [A1]: 默认 UMI/gene 质控过滤标准，如有调整可自行修改

批注 [A2]: 默认线粒体比例过滤标准，如有调整可自行修改

批注 [A3]: 质控最终得到的细胞数目

批注 [A4]: 如果进行了样本间去批次处理，可将这句话替换为：
To remove the batch effects in single-cell RNA-sequencing data, the mutual nearest neighbors(MNN) presented by Haghverdi et al was performed with the R package batchelor.

Haghverdi L, Lun A T L, Morgan M D, et al. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors[J]. Nature biotechnology, 2018, 36(5): 421-427.

identified positive markers compared with all other cells. Then, we used the R package SingleR[4], a novel computational method for unbiased cell type recognition of scRNA-seq, with the reference transcriptomic datasets 'Human Primary Cell Atlas' (Mabbott et al. 2013) to infer the cell of origin of each of the single cells independently and identify cell types.

Differentially expressed genes(DEGs) were identified using the FindMarkers function（test.use = MAST） in Seurat[1]. P value < 0.05 and $|\log_2 foldchange| > 0.58$ was set as the threshold for significantly differential expression. GO enrichment and KEGG pathway enrichment analysis of DEGs were respectively performed using R based on the hypergeometric distribution.

The sequencing and bioinformatics analysis were performed by OE Biotech Co., Ltd. (Shanghai, China).

[1] Butler A, Hoffman P, Smibert P, et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species[J]. Nature biotechnology, 2018, 36(5): 411-420.
[2] Mcginnis C S , Murrow L M , Gartner Z J . DoubletFinder: Doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. 2018.
[3] Macosko E Z, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets[J]. Cell, 2015, 161(5): 1202-1214.
[4] Aran D, Looney A P, Liu L, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage[J]. Nature immunology, 2019, 20(2): 163-172.

批注 [A5]: 细胞类型鉴定部分使用的参考数据集。可根据实际情况自行添加修改。

人的可选参考数据集：
1、HPCA（Human Primary Cell Atlas）：Mabbott N A, Baillie J K, Brown H, et al. An expression atlas of human primary cells: inference of gene function from coexpression networks[J]. BMC genomics, 2013, 14(1): 632.
2、blueprint+encode：
Blueprint:
Martens, J. H. A, Stunnenberg, H. G. BLUEPRINT: mapping human blood cell epigenomes[J]. Haematologica, 98(10):1487-1489.
Encode：
Bernstein B E, Birney E, Dunham I, et al. An integrated encyclopedia of DNA elements in the human genome[J]. Nature, 2012, 489(7414): 57-74.
3、schcl： http://bis.zju.edu.cn/HCL/index.html

小鼠可选参考数据集：
1、 immgen：Heng T S P , Painter M W , Elpek K , et al. The Immunological Genome Project: networks of gene expression in immune cells[J]. Nature Immunology, 2008, 9(10):1091-1094.
2、scmca：Xiaoping Han, Renying Wang, Yincong Zhou,et al. Mapping the Mouse Cell Atlas by Microwell-Seq[J]. Cell, 2018, 172(5):1091-1107.e17.
3、mouse.rnaseq：Benayoun B A, Pollina E A, Singh P P, et al. Remodeling of epigenome and transcriptome landscapes with aging in mice reveals widespread induction of inflammatory responses[J]. Genome research, 2019, 29(4): 697-709.