**Supplemental Methods**

**Tissue dissociation and preparation**

After filtering through 40-µm sterile strainers, ascites sample were centrifuged at 350 × *g* for 5 min. The supernatant was discarded, and the sediment was resuspended in 1 ml PBS (HyClone). GEXSCOPE® red blood cell lysis buﬀer (Singleron) was added, and the mixture [CellS: lysis buﬀer =1:2 (volume ratio)] was incubated at room temperature for 5–8 min to remove red blood cells. the samples were then stained with trypan blue, and cell viability was evaluated microscopically.

**RT, amplification, and library construction**

Single cell suspensions (2×105 cells/mL) in PBS (HyClone) were loaded onto a microwell chip using the Singleron Matrix® Single Cell Processing System. Barcoding beads were subsequently collected from the microwell chip, followed by reverse transcription of the mRNA captured by the barcoding beads and to obtain cDNA for PCR amplification. The amplified cDNA was fragmented and ligated to sequencing adapters. scRNA-seq libraries were constructed in accordance with the protocol of the GEXSCOPE® Single Cell RNA Library Kit (Singleron)1. Individual libraries were diluted to 4 nM, pooled, and sequenced on an Illumina novaseq 6000 with 150 bp paired-end reads.

**Primary analysis of raw read data**

Raw reads from scRNA-seq were processed to generate gene expression matrixes using the CeleScope v1.9.0 pipeline (https://github.com/singleron-RD/CeleScope). Briefly, raw reads were first processed with CeleScope to remove low quality reads with Cutadapt v1.172 to trim poly-A tail and adapter sequences. The cell barcode and UMI were extracted. We then used STAR v2.6.1a3 to map reads to the reference genome GRCh38 (ensembl version 92 annotation). UMI counts and gene counts of each cell were acquired with featureCounts v2.0.14 software and used to generate expression matrix files for subsequent analysis.

**Quality control, dimension reduction, and clustering**

Scanpy5 v1.8.2 was used for quality control, dimensionality reduction, and clustering under Python 3.7. For each sample dataset, we filtered the expression matrix by the following criteria: 1) cells with a gene count of less than 200 or with top the 2% gene count were excluded; 2) cells with the top 2% UMI count were excluded; 3) cells with >20% mitochondrial content were excluded; 4) genes expressed in less than five cells were excluded. After filtering, 17996 cells were retained for downstream analyses with an average of 869 genes and 2729 UMIs per cell. The raw count matrix was normalized by total counts per cell and logarithmically transformed into a normalized data matrix. The top 2000 variable genes were selected by setting flavor = ‘seurat’. Principle component analysis was performed on the scaled variable gene matrix, and the top 20 principle components were used for clustering and dimensional reduction. Cells were separated into 17 clusters using the Louvain algorithm and setting the resolution parameter at 1.2. Cell clusters were visualized using uniform manifold approximation and projection (UMAP) [t-distributed stochastic neighbor embedding (t-SNE)]. Batch effect between samples was removed by Harmony6 v1.0 using the top 20 principal components from principle component analysis.

**Statistics and repeatability**

Cell distribution comparisons between two groups were performed using the unpaired two-tailed Wilcoxon rank-sum test. Comparisons of gene expression or gene signatures between two groups of cells were performed using the unpaired two-tailed Student’s t-test. Comparisons of the cell distribution of paired group 1 and group 2 as well as after and before were performed using the paired two-tailed Wilcoxon rank-sum tests. All statistical analyses were performed using R. Statistical tests used in figures are shown in figure legends and statistical significance was set at p < 0.05. The exact value of n is shown in the figures and figure legends and what n represents is shown in the figure legends.

**Differentially expressed gene (DEG) analysis**

To identify differentially expressed genes (DEGs), we used the Seurat FindMarkers function based on the Wilcox likelihood-ratio test with default parameters, and selected genes expressed in > 10% of cells in a cluster with an average log (fold change) value greater than 0.25 as DEGs. For cell type annotation of each cluster, we used the expression of canonical marker DEGs and displayed the expression of markers of each cell type by heat maps, dot plots, or violin plots generated by Seurat DoHeatmap, DotPlot, and Vlnplot functions, respectively. Doublet cells were identified as expressing markers for different cell types and removed manually.

**Cell type annotation**

The cell type identity of each cluster was determined by the expression of canonical marker DEGs using the SynEcoSys database. Heat maps, dot plots, and violin plots displaying the expression of markers used to identify each cell type were generated by Seurat v3.1.2 DoHeatmap, DotPlot, and Vlnplot functions, respectively.

**Pathway enrichment analysis**

To investigate the potential functions of DEGs, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were used with “clusterProfiler” R package 4.0.27. Pathways with a p\_adj value less than 0.05 were considered to be significantly enriched. Gene Ontology gene sets, including molecular function, biological process, and cellular component categories, were used as references.

**Trajectory analysis**

Cell differentiation trajectory was reconstructed with Monocle28. Highly variable genes were used to sort cells in order of spatiotemporal differentiation. We used DDRTree to perform FindVairableFeatures and dimension-reduction. The trajectory was visualized by the plot\_cell\_trajectory function. CytoTRACE9 (a computational method that predicts the differentiation state of cells from single cell RNA-sequencing data using gene counts and expression) was used to predict the differentiation potential of monocyte subpopulations.

**RNA velocity**

For RNA velocity, BAM file containing pre-B cells, GMP, MPs, and reference genome GRCh38 (hg38) were used in the analysis with velocyto v0.17.17 and scVelo v0.2.4 in Python with default parameters10 11. The result was projected to the UMAP plot from Seurat clustering analysis for visualization consistency.

**scRNA-seq-based CNA detection**

The InferCNV package12 was used to detect CNAs in 5000 malignant cells. A total of 4203 non-malignant cells were used as baselines to estimate CNAs in malignant cells. Genes expressed in more than 20 cells were sorted by their loci on each chromosome. The relative expression values were centered to 1 using a 1.5 standard deviation from the residual-normalized expression values as the floor and ceiling. A slide window size of 101 genes was used to smoothen the relative expression on each chromosome to remove the effect of gene-specific expression. The inferred CNAs on each short or long arm, or full length chromosome were visualized using heatmaps generated by the R pheatmap function. The clonal relationships of malignant cell clusters in each sample were determined by CNA accumulation. UPhyloplot2 software was used to draw clonality trees of malignant cell clusters. CNA events were manually labeled on the branches. The CNV score of each cell was calculated as the quadratic sum of CNAregion12.

**Copy number variation (CNV) for whole exon sequencing analysis**

We used bwa alignment software13 (0.7.17) and human reference genome (hg38) to carry out sequence alignment on the sequencing data of whole exon sequencing. We used cnvkit software14 (0.9.9) to identify the CNV of the comparison results of paired samples. To calculate the CNV score of each gene, we used GISTIC2.0 software15. Finally, we used the ComplexHeatmap package (2.10.0) in R language to graphically display the CNV situation of each sample.

**Cell–cell interaction analysis**

Cell–cell interactions between monocytes, GMP, HSCs, and pre-B, T, NK, and B cells were predicted by known ligand–receptor pairs in CellphoneDB (v2.1.0)16. The permutation number to calculate the null distribution of average ligand–receptor pair expression in randomized cell identities was set to 1000. Individual ligand or receptor expression was thresholded by a cutoff based on the average log gene expression distribution for all genes across each cell type. Predicted interaction pairs with p < 0.05 and average log expression > 0.1 were considered to be significant and visualized by heatmap\_plot and dot\_plot in CellphoneDB.

1. Dura, B. *et al.* scFTD-seq: freeze-thaw lysis based, portable approach toward highly distributed single-cell 3' mRNA profiling. *Nucleic Acids Res* **47**, e16 (2019).

2. Kechin, A., Boyarskikh, U., Kel, A. & Filipenko, M. cutPrimers: A New Tool for Accurate Cutting of Primers from Reads of Targeted Next Generation Sequencing. *J Comput Biol* **24**, 1138-1143 (2017).

3. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

4. Liao, Y., Smyth, G.K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).

5. Wolf, F.A., Angerer, P. & Theis, F.J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**, 15 (2018).

6. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411-420 (2018).

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

8. Qiu, X. *et al.* Single-cell mRNA quantification and differential analysis with Census. *Nat Methods* **14**, 309-315 (2017).

9. Gulati, G.S. *et al.* Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405-411 (2020).

10. La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, 494-498 (2018).

11. Bergen, V., Lange, M., Peidli, S., Wolf, F.A. & Theis, F.J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* **38**, 1408-1414 (2020).

12. Tirosh, I. *et al.* Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* **539**, 309-313 (2016).

13. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

14. Talevich, E., Shain, A.H., Botton, T. & Bastian, B.C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol* **12**, e1004873 (2016).

15. Mermel, C.H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* **12**, R41 (2011).

16. Efremova, M., Vento-Tormo, M., Teichmann, S.A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* **15**, 1484-1506 (2020).