APR 15, 2024

Protocol Citation: Siwei Li 2024. Single-cell RNA-seq. **protocols.io** https://protocols.io/view/single-cell-rna-seq-db5h2q36

Protocol status: Working
We use this protocol and it's working

Created: Apr 15, 2024

Last Modified: Apr 15, 2024

PROTOCOL integer ID: 98185

Single-cell RNA-seq

Siwei Li[1]

[1]Harbin Medical University Cancer Hospital

Siwei Li
Harbin Medical University Cancer Hospital

ABSTRACT

Single-cell RNA-seq

Single-cell RNA-seq

1 Tissues were transferred to sterile culture dishes with 10 ml of 1× Dulbecco's phosphate-buffered saline (DPBS; Thermo Fisher, Cat. no. 14190144) on ice to remove residual tissue storage solution and then minced on ice. We used the dissociation enzyme 0.25% trypsin (Thermo Fisher, Cat. no. 25200-072) and 10 µg/mL lDNase I (Sigma, Cat. no. 11284932001) dissolved in PBS with 5% foetal bovine serum (FBS; Thermo Fisher, Cat. no. SV30087.02) to digest the tissues. XX tissues were dissociated at 37 °C with shaking at 50 rpm for approximately 40 min. We repeatedly collected the dissociated cells at intervals of 20 min to increase cell yield and viability. The cell suspensions were filtered through a 40 µm nylon cell strainer, and the red blood cells were removed with 1X Red Blood Cell Lysis Solution (Thermo Fisher, Cat. no. 00-4333-57). The dissociated cells were washed with 1× DPBS containing 2% FBS. The cells were stained with 0.4% Trypan blue (Thermo Fisher, Cat. no. 14190144) to assess viability on a Countess® II Automated Cell Counter (Thermo Fisher).

Beads with a unique molecular identifier (UMI) and cell barcodes were loaded close to saturation so that each cell was paired with a bead in a gel beads-in-emulsion (GEM) system. After exposure to cell lysis buffer, polyadenylated RNA molecules hybridized to the beads. The beads were transferred to a single tube for reverse transcription. For cDNA synthesis, each cDNA molecule was tagged on the 5’ end (that is, the 3’ end of a messenger RNA transcript) with UMI and a cell label indicating its cell of origin. Briefly, 10× beads were subjected to second-strand cDNA synthesis, adaptor ligation, and universal amplification. Sequencing libraries were prepared using randomly interrupted whole-transcriptome amplification products to enrich the 3’ end of the transcripts linked with the cell barcode and UMI. All the remaining procedures, including library construction, were performed according to the standard manufacturer’s protocol (Chromium Single Cell 3’ v3). The sequencing libraries were quantified using a high-sensitivity DNA chip (Agilent) on a Bioanalyzer 2100 and a Qubit high-sensitivity DNA assay (Thermo Fisher Scientific). The libraries were sequenced on a NovaSeq 6000 (Illumina) using 2x150 chemistry.

Reads were processed using the Cell Ranger 2.1.0 pipeline with the default and recommended parameters. FASTQs generated from Illumina sequencing output were aligned to the mouse genome, version GRCm38, using the STAR algorithm. Next, gene-barcode matrices were generated for each individual sample by counting UMIs and filtering non-cell-associated barcodes. Finally, we generated a gene-barcode matrix containing the barcoded cells and gene expression counts. This output was then imported into the Seurat (v2.3.0) R toolkit for quality control and downstream analysis of our single-cell RNA-seq data. All functions were run with default parameters unless specified otherwise. We excluded cells with fewer than 200 or more than 6000 detected genes (where each gene had to have at least one UMI aligned in at least three cells). The expression of mitochondrial genes was calculated using the PercentageFeatureSet function of the Seurat package. To remove low-activity cells, cells with more than 10% expression of mitochondrial genes were excluded. The normalized data (normalizeData function in the Seurat package) were used to extract a subset of variable genes. Variable genes were identified while controlling for the strong relationship between variability and average expression. Next, we integrated data from different samples after identifying ‘anchors’ between datasets using FindIntegrationAnchors and IntegrateData in the seurat package. Then, we performed principal component analysis (PCA) and reduced the data to the top 30 PCA components after scaling the data.

We visualized the clusters on a 2D map produced with t-distributed stochastic neighbour embedding (t-SNE).

Cells were clustered using graph-based clustering of the PCA reduced data with the Louvain method after computing a shared nearest neighbour graph. For subclustering, we applied the same procedure of scaling, dimensionality reduction, and clustering to the specific set of data (usually restricted to one type of cell). For each cluster, we used the Wilcoxon rank-sum test to find significant DEGs between the remaining clusters. SingleR and known marker genes were used to identify cell types.

Tissue samples from humans and mice were processed into paraffin sections and then subjected to mIHC. For mouse tissue, we utilized a multiplex fluorescence staining kit (abs50028, Absin). The experimental equipment, primary reagents and antibodies used can be found in Extended Data mIHC.