One-stop solution for single-cell and nanoscale resolution spatial transcriptome analysis

A brief introduction to MGI's scRNA-Seq and STOmics' Stereo-seq technique

This application note explains the technical route, experimental procedures and analysis methods of scRNA-Seq and Stereo-seq technology independently developed by MGI and STOmics. Moreover, several research studies based on those two techniques were annotated at the end of this document.

Recommended applications: scRNA-Seq, Stereo-Seq

Recommended models: DNBSEQ-G400, DNBSEQ-T7 (Sequencing platform)

- High throughput single-cell transcriptome analysis using MGI DNBelab C series High throughput (HT) scRNA library preparation set
- Large field of view and nanoscale resolution of spatial transcriptomics analysis using STOmics' Stereo-seq
- DNBSEQ-G400 and DNBSEQ-T7 sequencing platform offer medium-ultra high sequencing throughputs that matches the needs of single-cell and Stereo-seq analysis

Background

Single-cell RNA-sequencing (scRNA-seq) profiles gene expression at individual cell level and has become a powerful tool in identifying cellular heterogeneity, rare cell populations, cell lineage trajectories, and the molecular mechanisms underlying development, disease, and complex biological processes.

However, the scRNA-seq loses cellular spatial context information during the step of tissue dissociation, which limits the power of scRNA-seq in dissecting cell-to-cell interaction and spatial distribution of different cell types. Compared to single-cell transcriptomics, recently developed spatial transcriptomics sequencing technology preserves spatial information while profiling gene expression. The integration of spatial data into gene expression profiles transforms our understanding of tissue structure and cell-cell interactions, facilitating a multidimensional analysis of disease microenvironment[1,2].

MGI offers a range of products for scRNA-seq and STOmics provides tools for spatial transcriptomics analysis. When paired with MGI massively parallel sequencing (MPS) DNBSEQ systems, these tools enable the simultaneous profiling of tens of thousands of transcripts at a single-cell resolution with or without spatial context.

This application note offers an overview of the workflows for single-cell and spatial transcriptomics analysis using MGI's and STOmics' readily available library construction and sequencing products.

Workflow overview

DNBelab C series HT scRNA library preparation set, MGI's single-cell analysis product

The DNBelab C series HT scRNA library preparation set includes reagents for droplet generation, flow cells for sample loading, and library preparation reagents. This set collaborates seamlessly with the droplet generation device TaiM, supporting cell isolation and cell barcoding. The freely available C4-Tools data analysis package facilitates the data processing. Figure 1 is an overview of the analysis workflow.

STOmics Stereo-seq Transcriptomics Set, STOmics' spatial transcriptomics analysis product

For spatially resolved 3' mRNA library generation from biological tissue sections, the STOmics Spatial Enhanced Resolution Omics-Sequencing (Stereo-seq) Transcriptomics Set utilizes MGI's patented DNA Nanoball (DNB) technology[3]. This set enables comprehensive transcriptome analysis *in situ* to spatial context solutions through in situ capture of mRNA at nanoscale resolution (500 nm) and a centimeter-sized field of view. The Stereo-seq Transcriptomics Solution employs spatially barcoded probes to capture and prime p*in situ*lenylated mRNA from tissue sections in situ. Each cDNA synthesized from the mRNA captured at a specific spot is linked to its spatially barcoded probe, facilitating subsequent gene expression mapping of a tissue section. Spatially-resolved gene expression profile was then achieved through analysis and visualization of the sequencing data using the Stereo-seq Analysis Workflow (SAW) pipeline and visualization software - StereoMap. Figure 2 is an overview of the Stereo-seq transcriptomics workflow.

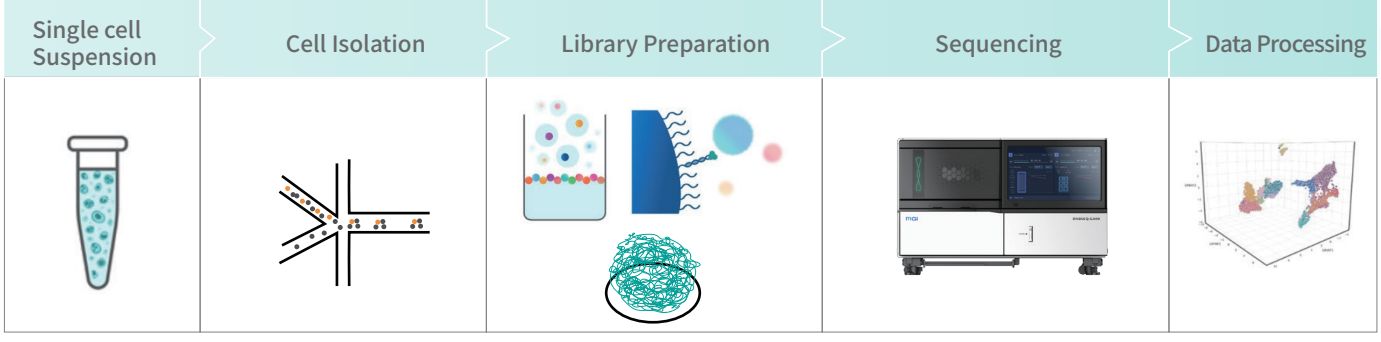

Figure 1. The workflow of DNBelab C series HT scRNA-seq analysis.

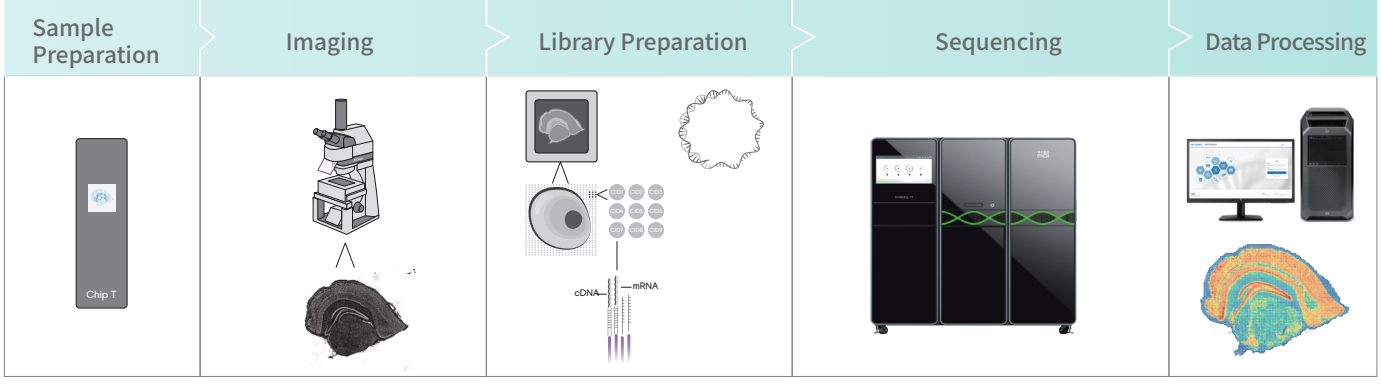

Figure 2. The workflow of Stereo-seq Transcriptomics analysis.

Materials and Methods

Sample preparation

For single-cell analysis, fresh tissue samples undergo enzymatic digestion, filtration and washing to obtain a single-cell suspension, with viability exceeding 80%. Nucleus suspension from fresh or frozen tissue can also be as input for single-nuclei analysis. It is essential to minimize the presence of dead cells and cellular aggregates in the cell and nuclei suspension to ensure good data quality. The recommended input for each sample is 20000 cells.

For Stereo-seq analysis, fresh tissue is first embedded in OCT within 30 min of resection. After freezing using dry ice, embedded samples are transferred to an -80 ℃ freezer for storage until use. Then, sections (10 μm) were cut from fresh frozen tissue blocks using a cryostat and promptly affixed to Stereo-seq chips for library construction.

DNBelab C series HT scRNA library preparation set is compatible with freshly isolated cells, or cryopreserved cells as well as isolated nuclei. The STOmics Stereo-seq Transcriptomics Set is compatible with fresh frozen and fixed frozen samples. Both DNBelab C series HT scRNA and Stereo-seq have been validated and tested in different tissue types from various species, including human, mouse, plants and insects.

Library preparation

Single-cell RNA sequencing libraries are prepared using the DNBelab C4 scRNA Preparation Kit according to the user manual, involving dilution of cells, microfluidic chip loading, droplet generation, mRNA capture, reverse transcription, PCR amplification, cDNA purification and subsequent library construction with size selection and adapter ligation. The final libraries are purified and quantified before sequencing. Relevant instructional manuals could be referred to for detailed operational procedures.

Spatial transcriptomic data are generated using the STOmics' Stereo-seq Transcriptomics Set and the Chip-on-a-slide protocol is available at https://en.stomics.tech/resources/documents/list.html.

Tissue sections were adhered to the Stereo-seq chip, stained and imaged. RNA capture on the chip, reverse transcription, cDNA release and library construction were performed, including steps such as fragmentation, amplification and purification.

Sequencing

Above mentioned scRNA-seq libraries and Stereo-seq libraries are compatible with all DNBSEQ sequencers. However, higher throughput instruments such as the DNBSEQ-G400, and DNBSEQ-T7 are recommended to meet the required sequencing depth. The required sequencing amount will vary depending on the input cell number and the sample types. For scRNA-seq, The recommended sequencing amount is 600-800 million reads targeting 10000 cells. For Stereo-seq, 2-3 billion reads are recommended for a 1cm × 1cm capture area.

The scRNA-seq cDNA library was sequenced with the following sequencing length: 30 base pairs (bp) for read 1, including 20 bp cell barcode

and 10 bp unique molecular identifier (UMI), 100 bp of transcript sequence for read 2, and 10 bp for the sample index. For Stereo-seq library, the sequencing length is 35 bp for read 1, including 25 bp coordinate identity (CID) and 10 bp unique UMI, 100 bp for read 2 and 10 bp for the sample index.

Data analysis

The scRNA-seq raw reads FASTQ were processed using the MGI C4 tool software package (https://github.com/MGI-tech-bioinformatics/DNBelab_C_Series_HT_scRNA-analysis-software). The initial step of this preprocessing procedure is to clean and filter the scRNA-seq data in order to get rid of artifacts like contaminants, low-quality cells, and doublets. Additionally, it ensured a sufficient number of detected genes in the remaining cells. After the UMI normalization and log transformation, the data are ready for downstream analysis.

For Stereo-seq data analysis and visualization, MGI and STOmics provide a variety of data analysis methods. Apart from the STOmics Cloud, STOmics offers a comprehensive offline software suite consisting of ImageStudio, SAW, and StereoMap software. Stereo-seq FASTQ files from the DNBSEQ-T7 sequencer underwent the following process. Reads with MID quality score <10 were eliminated, CID sequences were mapped to chip coordinates, and cDNA sequences were aligned to the reference genome using the SAW program. Data matrix files were generated, and for sub-cellular resolution, DNBs were merged into a 'bin' for downstream analysis.

Application highlights

Hight throughput scRNA-seq to generate cell transcriptomic atlas of the non-human primate Macaca fascicularis

Using DNBelab C4 scRNA library preparation system and DNBSEQ sequencing platforms, a large-scale single-cell transcriptome analysis was performed on the main tissues and organs (lung, kidney, liver, pancreas, brain, aorta, thyroid, parotid gland, and blood) of cynomolgus monkeys. This effort constructed the first non-human primate single-cell atlas that consists of 1 million cells across 45 tissues (Figure 3)[4].

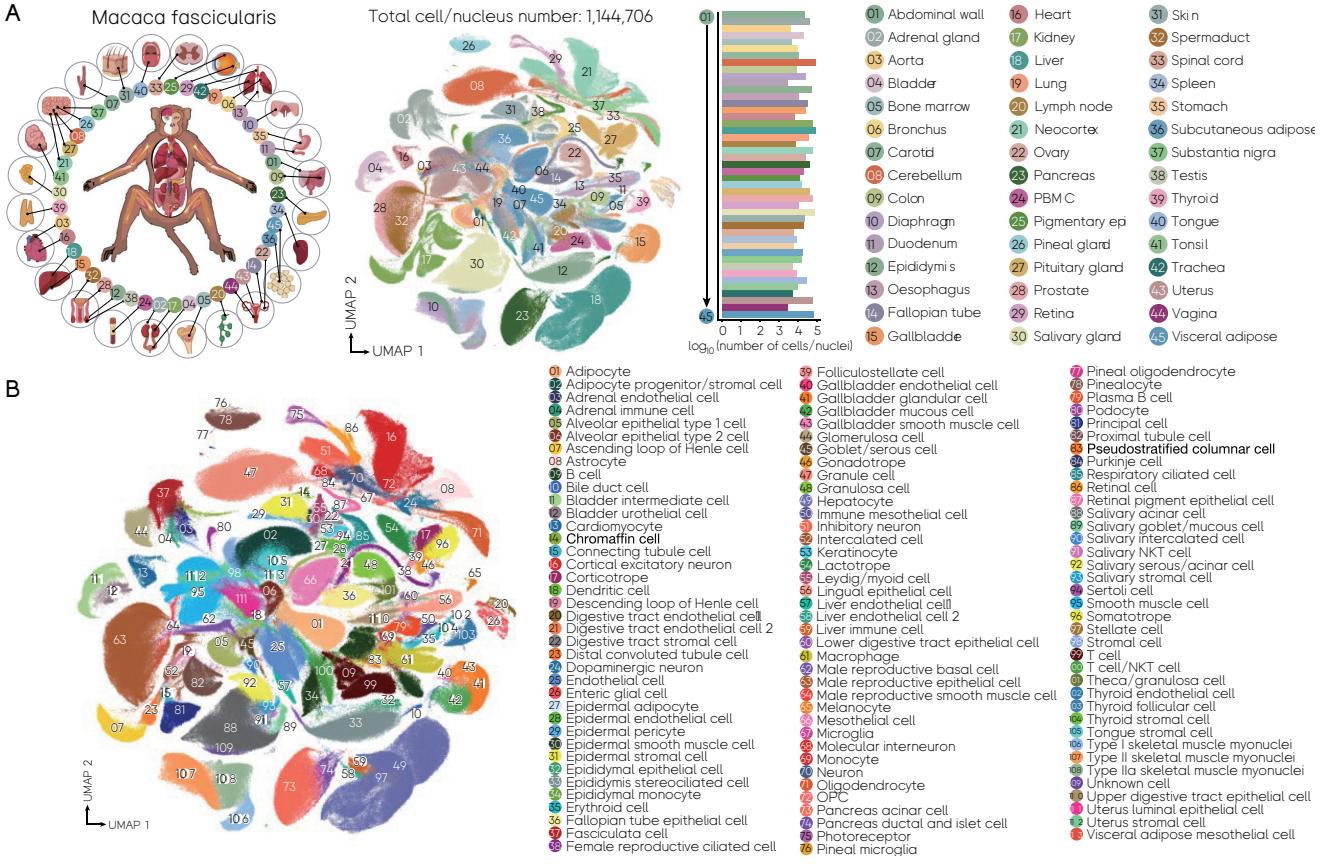

Figure 3. **Identification of different cell types across 45 tissues of adult M. fascicularis monkey.**
(A) On the left is a Schematic illustration of the monkey tissues examined in this investigation, in total 45 tissues were taken from 6-year-old monkeys, 3 of whom were female and 5 of whom were male. UMAP (middle) presents the global clustering of all the cells in the dataset, colored by tissue, and bar plots (right) display the number of cells/nuclei profiled for each tissue following quality control. 1,144,706 distinct cells or nuclei were examined (n = 1). (B) A UMAP visualization showing the primary cell types that color each cluster. In the dataset, 113 distinct cell clusters were found. The right lengend provides cell type annotation for each main cluster.

Spatial heterogeneity of cell types in mouse embryonic tissues

Spatial transcriptomic analysis of embryos is critical to understanding the dynamics of cell fate during tissue development. Using Stereo-seq and DNBSEQ sequencing platform, gene expression with spatial context of 281,377 cells (an average of 1,107 MIDs and 529 genes per cell) was obtained from E16.5 mouse embryos. Comprehensively describing the different cell types and their spatial relationships at the level of the entail embryo is important for understanding the mechanisms involved in cell differentiation, and regional specificities (Figure 4)[3].

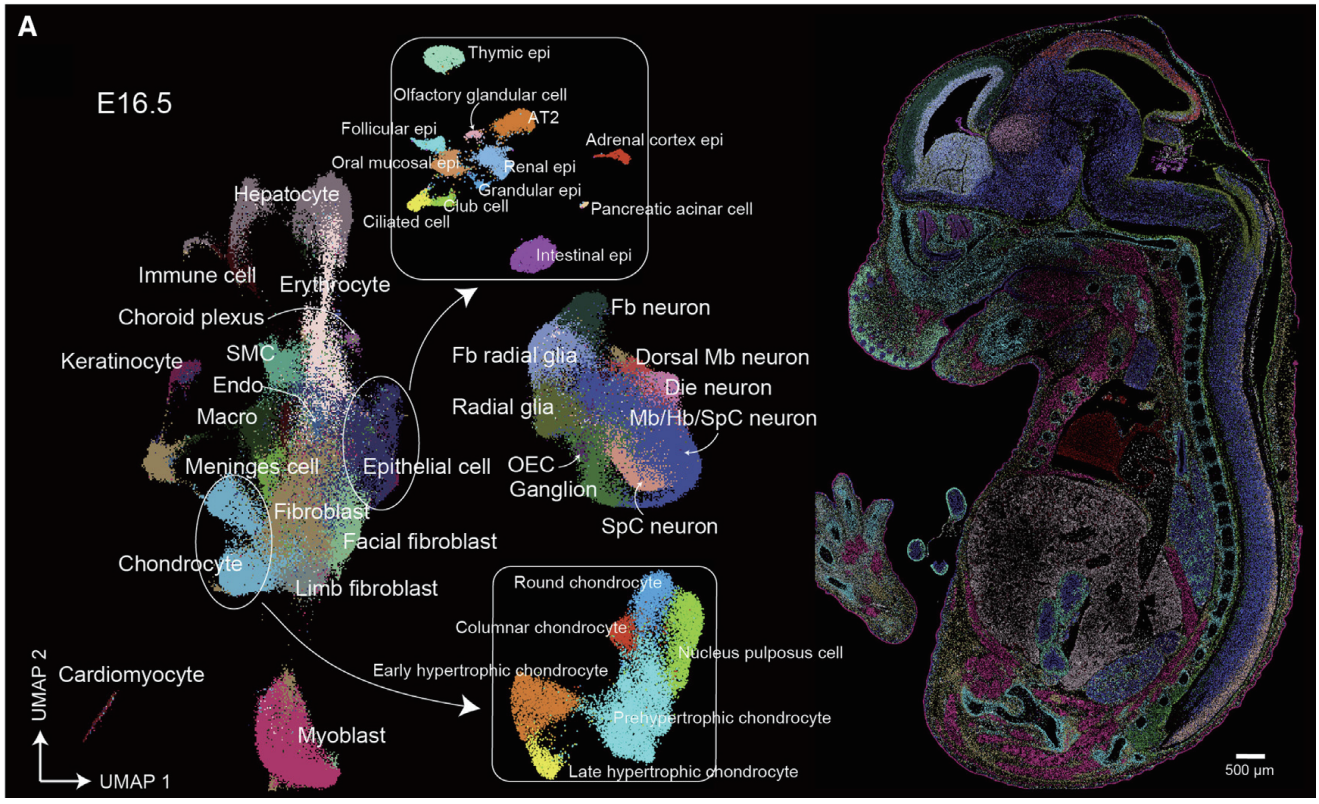

Figure 4. **Spatial diversification of cell types at whole embryo scale**.
(A) Segmented cells from an E16.5 slice (E1S3) with UMAP on the left. The side panels provide the UMAP visualization of the chondrocyte (bottom) and epithelial cell (top) re-clustering outcomes. Right: a spatial representation of the cell types for the whole E16.5 embryo segment, as seen in the left panel. Annotation colorizes the cells. Fb neuron, forebrain neuron; Dorsal Mb neuron, dorsal midbrain neuron; Die neuron, diencephalon neuron; Mb/Hb/SpC neuron, mid-/hindbrain, and spinal cord neuron; SMC, smooth muscle cell; Endo, endothelial cell; and Macro, macrophage. The scale bars, 500 μm.

Dynamics of cell composition and tumor microenvironment following treatment revealed by scRNA-seq and Stereo-seq in rectal cancer

To examine the dynamics of the tumor microenvironment (TME) in rectal cancer (RC) following NAC treatment, 29 patients were recruited, and paired tumor endoscopic biopsies and surgical resections were collected pretherapy (PE) and post-therapy (PO), respectively. Based on the tumor regression grade (TRG) evaluation of the post-therapy sample, patients were divided into three groups, complete (or nearly complete) regression (CR), partial regression (PR), and no regression (NR). Analysis of the samples before and after treatment using an integrated approach of scRNA-seq and Stereo-seq revealed cellular dynamic landscape and TME remodeled by treatment and their associations with therapeutic response (Figure 5)[5].

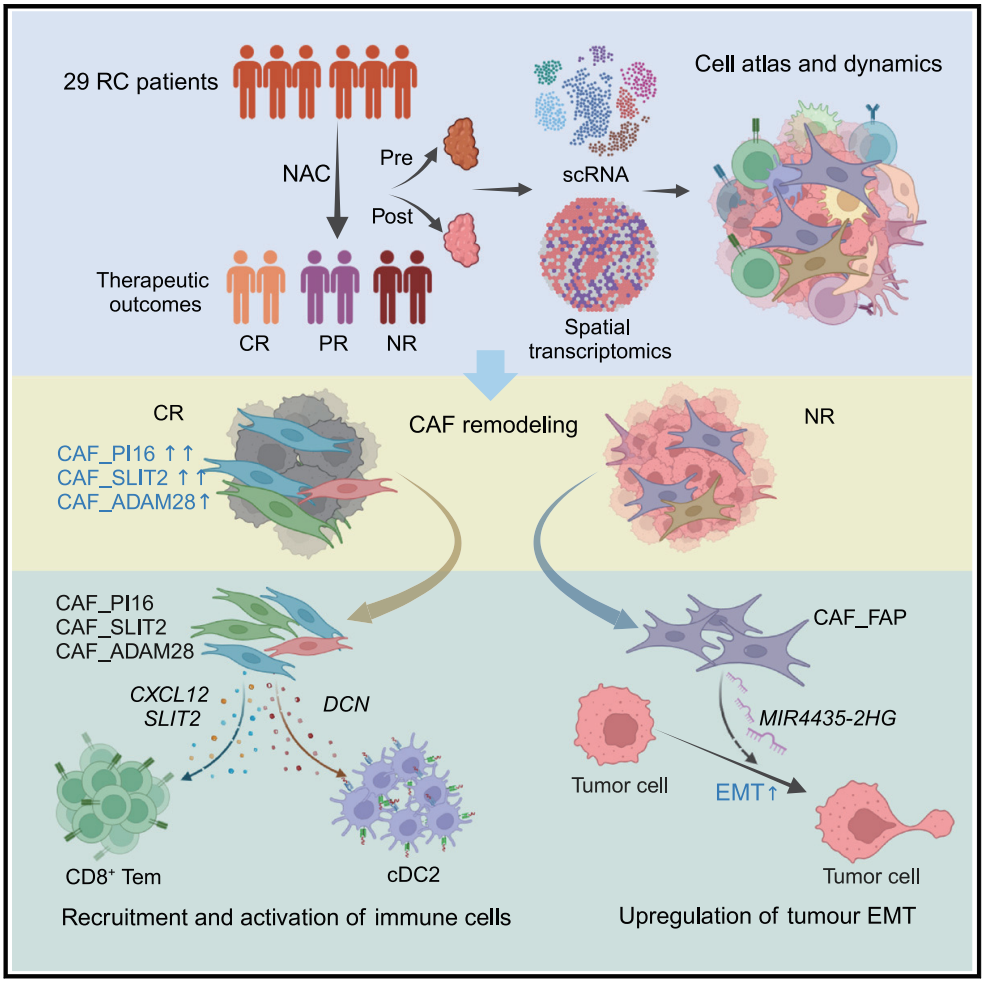

Figure 5. **The experiment design and major findings of the study investigating** the modulation of the tumor microenvironment (TME) by NAC and its association with the therapeutic response using an integrated approach of single-cell and Stereo-seq. RC: rectal cancer, NAC: Neoadjuvant chemotherapy , CR: complete (or nearly complete) regression, PR: partial regression, NR: no regression (NR).

Summary

This application note offers a brief overview of workflows for scRNA-Seq and Stereo-seq based on MGI's DNBSEQ sequencing platform. These technologies empower researchers to comprehensively investigate cellular diversity, cellular heterogeneity, and spatial organization of different cell types, to gain a better understanding of biological processes that underly health and disease.

DNBSEQ-G400 Genetic Sequencer

DNBSEQ-T7 Genetic Sequencer

References

1. Saviano, A., Henderson, N. C. & Baumert, T. F. Single-cell genomics and spatial transcriptomics: Discovery of novel cell states and cellular interactions in liver physiology and disease biology. *Journal of Hepatology* **73**,1219-1230,doi:10.1016/j.jhep.2020.06.004 (2020).
2. Zheng, B. & Fang, L. Spatially resolved transcriptomics provide a new method for cancer research. *Journal of Experimental & Clinical Cancer Research* **41**, doi:10.1186/s13046-022-02385-3 (2022).
3. Chen, A. *et al*. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777-1792.e1721,doi:10.1016/j.cell.2022.04.003 (2022).
4. Han, L. *et al*. Cell transcriptomic atlas of the non-human primate Macaca fascicularis. *Nature* **604**, 723-731,doi:10.1038/s41586-022-04587-3 (2022).
5. Qin, P. *et al*. Cancer-associated fibroblasts undergoing neoadjuvant chemotherapy suppress rectal cancer revealed by single-cell and spatial transcriptomics. *Cell Reports Medicine* **4**, doi:10.1016/j.xcrm.2023.101231 (2023).

Recommended Ordering Information

Application	Category	Product	Cat. NO.
Single cell RNA-seq	Instruments	Single Cell Droplet Generator DNBelab C-TaiM 4RS	900-000780-00
	Library Prep Reagents	DNBelab C Series High-throughput Single-cell RNA Library Preparation Set V2.0	940-000519-00
Spatio-temporal Transcrip-tomic	Library Prep Reagents	Stereo-seq Permeabilization Set for Chip-on-a-slide (8 RXN)	211SP118*
		Stereo-seq Transcriptomics Set for Chip-on-a-slide (4 RXN)	211ST114*
		Stereo-seq Library Preparation Kit (4 RXN)	111KL114*
	Sequencing Reagents	DNBSEQ Onestep DNB Make Reagent Kit (0s-SB)	1000020563
		High-Throughput Sequencing Primer Kit (Stere Omics)	940-000037-00
DNBSEQ Platform	Instrurnent	Geneic Sequencer DNBSEQ-G400RS	900-000170-00
		Genetic Seguencer DNBSEQ-T7RS	900-0002 42-00
	Sequencing Reagents	DNBSEQ-G400RS High-throughput Sequencing Set (SM FCL PE100)	100001 6950
		DNBSEQ-T7RS High-throughput Sequencing Set (SM FCL PE100) V2.0	1000028455

*The relevant products are available on the STOmics official website https://en.stomics.tech/.
For kits and reagents ordering, please consult info_global@stomics.tech

MGI Tech Co.,Ltd

Building 11, Beishan Industrial Zone, Yantian District, Shenzhen, CHINA, 518083

Version: February 2024

+86-4000-688-114

en.mgi-tech.com

MGI-service@mgi-tech.com

Authors: Hui Gao

Editor-in-Charge: Qiwei Wang

Reviewer: Yao Jiang