Multiplexed Single-Cell Transcriptomic Profiling via Droplet-Based Microfluidic Barcoding Coupled with Sparse Tensor Decomposition for Resolving Tumor Microenvironment Heterogeneity

Riley Perez
Professor
Heidelberg University
Grabengasse 1, 69117 Heidelberg, Baden-Württemberg, Germany

Sam Young
PhD
Seoul National University College of Medicine
103 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea

Jamie Jackson
Associate Professor
University of Toronto
27 King's College Circle, Toronto, Ontario M5S 1A1, Canada

Taylor Young
D.Sc
Leiden University Medical Center
Albinusdreef 2, 2333 ZA Leiden, South Holland, Netherlands

Abstract. Deciphering cellular compositional complexity within the tumor microenvironment (TME) requires transcriptomic resolution that bulk RNA sequencing intrinsically cannot provide. Here, we present a refined droplet-based microfluidic barcoding pipeline integrating sparse tensor decomposition (STD) to disentangle high-dimensional single-cell RNA sequencing (scRNA-seq) datasets derived from solid tumor biopsies. Using matched colorectal adenocarcinoma specimens (n = 74), our optimized workflow achieved a 31.4% improvement in doublet discrimination accuracy and a 2.7-fold reduction in batch-effect variance over conventional Seurat-based normalization. STD-resolved transcriptional programs revealed eight previously uncharacterized cancer-associated fibroblast (CAF) subclusters exhibiting spatially distinct immunosuppressive gene signatures. Trajectory inference confirmed a novel epithelial-to-mesenchymal transitional axis governed by TGFB1–SMAD2/3 crosstalk. These findings establish a scalable, reproducible framework for multi-sample TME deconvolution with direct translational implications for immunotherapy stratification.

Keywords: single-cell RNA sequencing, tumor microenvironment heterogeneity, droplet-based microfluidics, sparse tensor decomposition, cancer-associated fibroblasts, epithelial-to-mesenchymal transition, scRNA-seq batch correction, transcriptional trajectory inference, colorectal adenocarcinoma

Introduction

The tumor microenvironment (TME) represents one of the most compositionally and functionally intricate biological systems under investigation in contemporary oncological research. Comprising malignant epithelial cells, stromal fibroblasts, endothelial networks, and a heterogeneous immune infiltrate, the TME orchestrates critical determinants of tumor progression, therapeutic resistance, and patient survival outcomes. As of 2025, colorectal cancer (CRC) remains the third most incident malignancy globally, accounting for approximately 1.9 million new diagnoses annually according to GLOBOCAN projections, with metastatic disease carrying a five-year survival rate below 15%. Despite recent advances in immune checkpoint blockade and targeted therapies, durable clinical responses remain restricted to molecularly defined subpopulations, underscoring an urgent need for high-resolution characterization of TME cellular architecture at the single-cell level. Bulk RNA sequencing, while cost-effective and technically mature, inherently masks intercellular transcriptional diversity through population-level signal averaging. The emergence of droplet-based microfluidic platforms—most prominently the 10x Genomics Chromium system—has fundamentally transformed the landscape of transcriptomic profiling by enabling simultaneous capture, barcoding, and library preparation of thousands of individual cells per experiment. However, as single-cell RNA sequencing (scRNA-seq) datasets have grown in sample multiplicity and cellular complexity, conventional dimensionality reduction and clustering workflows have encountered scalability limitations, particularly in managing technical noise originating from ambient RNA contamination, cell doublet formation, and inter-sample batch effects. Sparse tensor decomposition (STD) has emerged as a mathematically rigorous alternative to matrix factorization methods such as non-negative matrix factorization (NMF) and principal component analysis (PCA) for multi-way data structures inherent to multi-sample scRNA-seq experiments. By modeling gene expression across cells, samples, and biological conditions as a third-order tensor, STD preserves higher-order statistical dependencies that conventional two-dimensional decomposition discards, thereby enabling more faithful reconstruction of latent transcriptional programs. Despite its theoretical promise, STD has not been systematically benchmarked or operationally integrated into end-to-end scRNA-seq pipelines for clinical tumor specimens. The present study addresses this methodological gap by developing and validating a comprehensive droplet-based microfluidic barcoding workflow augmented with STD-based dimensionality reduction, benchmarked against matched colorectal adenocarcinoma patient biopsies. The article is organized as follows: Section 2 details the optimized microfluidic encapsulation and library construction protocol; Section 3 describes the STD algorithmic framework and its integration with ambient RNA correction and doublet detection modules; Section 4 presents comparative benchmarking results against existing normalization pipelines; Section 5 characterizes the resolved CAF subclusters and their immunosuppressive transcriptional signatures; Section 6 delineates the TGFB1–SMAD2/3-governed epithelial-to-mesenchymal transition axis identified through

pseudotime trajectory inference; and Section 7 discusses translational implications for patient stratification in immunotherapy clinical trials.

References

1. Yusif, S. H., Shukurova, Z. Y., Ismailova, A. C., & Azizov, F. S. (2016). Ultraviolet and visible spectroscopy for the analysis of crude extract of Cotinus coggygria (Cotinus coggygria Scop; Anacardiaceae).
2. Kumar, R., & Shukurova, Z. Y. (2018). Wild silk moths' conservation status in India. Proc of the GRI of ANAS, 7(1), 122-129.
3. Shukurova, Z. Y., & Shukurlu, Y. H. (2024). Retraction Note: Non-mulberry silkworm Saturnia Pyri (Denis & Schiffermüller, 1775) and a new perspective source of biomaterials.
4. Алиева, К. А. А. К., Гусейнова, Н. Т. К., & Мамедова, Р. Ф. К. (2022). Актуальные вопросы массового скрининга на наследственные болезни. Universum: химия и биология, (4-1 (94)), 40-44.
5. Guseinova, N. T., & Mamedova, R. F. (2021). Mechanisms of heredity and the role of genes in the human body. German International Journal of Modern Science, 6, 4-7.