

Анализ данных транскриптомики единичных клеток на примере онкологических заболеваний

Иконников А.В.^{1, 4*}, Волчков Е.В.^{1, 2}, Хозяинова А.А.³, Меньяло М.Е.^{1, 3},
Копанцева Е.Е.¹, Гуржиханова М.Х.², Порядина Е.А.¹, Фетисов Т.И.⁵,
Якубовская М.Г.⁵, Денисов Е.В.^{1, 3}, Кирсанов К.И.⁵, Масчан М.А.²

¹ НИИ молекулярной и клеточной медицины МИ РУДН, Москва, Россия

² НМИЦ детской гематологии, онкологии и иммунологии им. Д. Рогачева, Москва, Россия

³ Томский НИМЦ РАН, Томск, Россия

⁴ Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, Россия

⁵ ФГБУ «НМИЦ онкологии им. Н.Н. Блохина» Минздрава России, Москва, Россия

* alex.v.ikonnikov@gmail.com

Ключевые слова: биоинформатика; single-cell RNA-seq; рак; ЮММЛ; саркома

Мотивация и цель: Анализ данных секвенирования РНК одиночных клеток (scRNA-seq) позволяет получать информацию о профилях экспрессии опухолевых клеточных популяций и изучать нарушения молекулярных механизмов на уровне отдельных клеток. Понимание клеточной гетерогенности и динамики клеточных взаимодействий в опухоли может привести к разработке новых комбинированных терапевтических стратегий, направленных на более эффективную терапию онкологических заболеваний [1–3].

Целью данного исследования являлся анализ данных РНК-секвенирования единичных клеток образцов пациентов с недифференцированными плеоморфными саркомами и образцов костного мозга (КМ) пациентов с ювенильным миеломоноцитарным лейкозом (ЮММЛ).

Методы и алгоритмы: В качестве образцов группы ЮММЛ были использованы клетки КМ 3 пациентов с мутациями генов RPTN11, NF11, NRAS. Контрольной группой являлись образцы scRNA-секвенирования КМ 3 здоровых детей, полученные из открытой базы данных NCBI GEO. Также было получено 9 образцов клеток тканей пациентов с недифференцированными плеоморфными саркомами. Протокол подготовки для РНК-секвенирования единичных клеток выполнялся на платформе Chromium X (10x Genomics, США). Секвенирование проводилось с помощью Genolab M (GeneMind, Китай). Биоинформатический анализ, включающий фильтрацию клеток и генов, нормализацию данных, интеграцию и типирование клеточных типов, выполнялся с помощью инструментов Scanpy (v.1.9.5) [4] и Seurat (v.5.0.3) [5]. Коррекция батч-эффекта после интеграции образцов было выполнено с помощью пакета Harmony [6]. Анализ метаболических путей выполнялся с использованием пакетов GSEApy [7] и DecoupleR [8].

Результаты: Обнаружено, что образцы ЮММЛ отличались повышенной пролиферацией клеток миеломоноцитарного ряда в сравнении с контролем. Образец ЮММЛ с агрессивным клиническим течением и мутацией в RPTN11 отличался значительным количеством стволовых клеток (HSCs), ранних предшественников и клеток эритроидного ряда. Анализ метаболических путей выявил, что клетки образцов ЮММЛ характеризовались повышенной активностью метаболического пути MAPK в сравнении с контролем. Данная

активность была неравномерна в клеточных популяциях и наиболее выражена в HSCs и MEPs (megakaryocyteerythroid progenitor cells). Также была обнаружена гиперактивация Р1ЗК-пути в образце ЮММЛ с мутацией *PTPN11*. Выявлено преобладание активности путей, связанных с ГТФазами и деацетилированием гистонов в клетках GMPs (granulocyte/monocyte progenitors), в образцах ЮММЛ в сравнении с контролем.

С помощью биоинформатической аннотации был охарактеризован клеточный ландшафт образцов ткани недифференцированных плеоморфных сарком. Обнаружены кластеры опухоль-ассоциированных фибробластов (CARs), опухолевых клеток, макрофагов, Т клеток и минорных редких клеточных популяций, характерные для отдельных пациентов.

Выводы: В результате исследования были изучены профили экспрессии генов различных клеточных популяций образцов КМ ЮММЛ в сравнении с образцами КМ контрольной группы. Были идентифицированы клеточные типы, присутствующие в образцах недифференцированных плеоморфных сарком.

Финансирование: Работа выполнена при финансовой поддержке Российского научного фонда (проект № 23-65-00003).

Analysis of single-cell transcriptomics data using the example of cancer diseases

Ikonnikov A.V.^{1,4*}, Volchkov E.V.^{1,2}, Khozyainova A.A.³, Menyailo M.E.^{1,3}, Kopantseva E.E.¹, Gurzhikhanova M.Kh.², Poryadina E.A.¹, Fetisov T.I.⁵, Yakubovskaya M.G.⁵, Denisov E.V.^{1,3}, Kirsanov K.I.⁵, Maschan M.A.²

¹ Research Institute of Molecular and Cellular Medicine, Peoples' Friendship University of Russia (RUDN University), Moscow, Russia

² Dmitry Rogachev National Research Center of Pediatric Hematology, Oncology and Immunology, Moscow, Russia

³ Tomsk National Research Medical Center (TNRMC), Tomsk, Russia

⁴ Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, Russia

⁵ Federal State Budgetary Institution, "N.N. Blokhin National Medical Research Center of Oncology" of the Ministry of Health of the Russian Federation (N.N. Blokhin NMRCO), Moscow, Russia

* alex.v.ikonnikov@gmail.com

Key words: bioinformatics; single-cell RNA-seq; cancer; JMML; sarcoma

Motivation and Aim: Single-cell RNA sequencing (scRNA-seq) analysis allows for the characterization of expression profiles of tumor cell populations and the study of disruptions in molecular mechanisms at the level of individual cells. Understanding cellular heterogeneity and dynamics of cell-cell interactions in tumors may lead to the development of novel combined therapeutic strategies aimed at more effective treatment of oncological diseases [1–3].

The aim of this study was to analyze single-cell RNA-seq data from samples obtained from patients with undifferentiated pleomorphic sarcomas and bone marrow (BM) samples obtained from patients with juvenile myelomonocytic leukaemia (JMML).

Methods and Algorithms: Bone marrow cells from 3 patients with mutations in the *PTPN11*, *NF11* and *NRAS* genes were used as samples from the JMML group. The

control group consisted of bone marrow samples obtained by scRNA sequencing from 3 healthy children obtained from the NCBI GEO open database. 9 tissue cell samples were also obtained from patients with undifferentiated pleomorphic sarcomas. The preparation of samples for single-cell RNA analysis was carried out on the Chromium X platform (10x Genomics, USA). Sequencing was performed using Geolab M (GeneMind, China). Bioinformatic analysis, including cell and gene filtering, data normalization, integration and cell type annotation, was performed using the Scanpy (version 1.9.5) [4] and Seurat (version 5.0.3) [5] packages. Correction of the batch effect after sample integration was performed using the Harmony [6] package. The analysis of metabolic pathways was conducted using the GSEAPy [7] and DecoupleR [8] packages. **Results:** It was found that JMML samples were characterized by increased proliferation of myelomonocytic cells compared to the control. A JMML sample with an aggressive clinical course and a mutation in PTPN11 was characterized by significant numbers of stem cells (HSCs), early progenitors, and erythroid cells. Analysis of metabolic pathways revealed that cells from JMML samples were characterized by increased activity of the MAPK metabolic pathway compared to controls. This activity was uneven across cell populations and was most pronounced in HSCs and MEPs (megakaryocyteerythroid progenitor cells). Hyperactivation of the PI3K pathway was also detected in a JMML sample with a PTPN11 mutation. The predominance of the activity of pathways associated with GTPases and histone deacetylation in GMPs (granulocyte/monocyte progenitors) cells in JMML samples was revealed in comparison with the control.

As a result of bioinformatics annotation, the cellular landscape of undifferentiated pleomorphic sarcoma samples was characterized. Clusters of tumor-associated fibroblasts (CARs), tumor cells, macrophages, T cells and minor rare cell population characteristic of individual patients were detected.

Conclusion: The study analyzed the gene expression profiles of various cell populations from JMML BM samples, comparing them with healthy control BM samples. The cellular types present in samples of undifferentiated pleomorphic sarcomas were identified.

Funding: The study is supported by the Russian Science Foundation (project No. 23-65-00003).

Список литературы/References

1. Chen S. et al. Single-cell analysis technologies for cancer research: from tumor-specific single cell discovery to cancer therapy. *Front Genet.* 2023;14:1276959
2. Louka E. et al. Heterogeneous disease-propagating stem cells in juvenile myelomonocytic leukemia. *J Exp Med.* 2021;218(2):e20180853
3. Yuan L.L. et al. Single-cell sequencing reveals the landscape of the tumor microenvironment in a skeletal undifferentiated pleomorphic sarcoma patient. *Front Immunol.* 2022;13:1019870. doi 10.3389/fimmu.2022.1019870
4. Wolf F.A. et al. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* 2018;19(1):15. doi 10.1186/s13059-017-1382-0
5. Hao Y. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol.* 2024;42(2):293-304
6. Korsunsky I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods.* 2019;16(12):1289-1296
7. Fang Z. et al. GSEAPy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics.* 2023;39(1):btac757
8. Badia-I-Mompel P. et al. decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinform Adv.* 2022;2(1):vbac016. doi 10.1093/bioadv/vbac016