

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

имени М. В. ЛОМОНОСОВА

*На правах рукописи*

**Арбатский Михаил Спартакович**

**Выяснение механизмов развития гетерогенного ответа мезенхимных стромальных клеток на профибротические стимулы с использованием анализа транскриптома единичных клеток**

Специальность 1.5.8 Математическая биология, биоинформатика

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени

кандидата биологических наук

Москва - 2026

Работа выполнена на кафедре биохимии и регенеративной биомедицины Факультета фундаментальной медицины Медицинского научно-образовательного института Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова».

Научный руководитель: ***Ефименко Анастасия Юрьевна***  
*доктор медицинских наук*

Официальные оппоненты: ***Колпаков Федор Анатольевич***  
*доктор биологических наук, Научно-технологический университет Сириус, Научный центр генетики и наук о жизни, направление «Вычислительная биология», научный руководитель*

***Храмеева Екатерина Евгеньевна***  
*доктор биологических наук, Сколковский институт науки и технологий, Центр биомедицинских технологий, доцент*

***Шеваль Евгений Валерьевич***  
*доктор биологических наук, Московский государственный университет имени М.В. Ломоносова, Научно-исследовательский институт физико-химической биологии имени А.Н.Белозерского, отдел электронной микроскопии, лаборатория ультраструктуры клеточного ядра, заведующий лабораторией*

Защита диссертации состоится 4 марта 2026 года в 17:30 на заседании диссертационного совета МГУ.015.10 Московского государственного университета имени М.В. Ломоносова по адресу: 119234, Москва, Ленинские горы, д. 1, стр. 73, Факультет биоинженерии и биоинформатики, ауд. 221.

E-mail: [dissovet@belozersky.msu.ru](mailto:dissovet@belozersky.msu.ru)

С диссертацией можно ознакомиться в отделе диссертаций Научной библиотеки МГУ имени М.В. Ломоносова (Москва, Ломоносовский просп., д. 27) и на портале:

<https://dissovet.msu.ru/dissertation/3611>

Автореферат разослан «\_\_» января 2026 года.

Ученый секретарь диссертационного совета,  
доктор биологических наук

 Д.В. Чистяков

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность и степень разработанности темы исследования.** С момента появления принципиально новых возможностей в анализе биологических образцов на уровне единичных клеток начало развиваться новое направление в биоинформатике – анализ данных РНК секвенирования единичных клеток (scRNA-seq). На текущий момент сформирована необходимая инфраструктура для этого направления, включая наборы реактивов, платформы для подготовки библиотек, библиотеки пакетов и программные конвейеры для процессинга данных. Особый интерес в изучении нового направления в биоинформатике представляет постоянное развитие методов обработки данных scRNA-seq. После освоения основных возможностей новой технологии продолжают появляться принципиально новые подходы к анализу данных scRNA-seq. Это связано с существенно более сложной структурой и объемом данных.

Разработка новой технологии была продиктована актуальной проблемой определения типа отдельных клеток или их субпопуляций, а также подробного описания молекулярных процессов в отдельных клетках. Внедрение scRNA-seq позволило получить ценную информацию о новых субпопуляциях клеток, присутствие которых в тканях определяет патогенез таких актуальных для здравоохранения состояний, как онкологические заболевания, метаболические нарушения и фибротические изменения органов и тканей.

Способность тканей млекопитающих, в том числе человека, восстанавливать свою структуру и функции после повреждения обеспечивается за счет репаративной регенерации, однако при длящемся или обширном повреждении, сопровождающемся развитием хронического воспаления, исходом этого процесса часто является возникновение фиброза. При этом происходит избыточное отложение белков внеклеточного матрикса (ВКМ) и его ремоделирование, что приводит к замещению функциональной ткани соединительной и в конечном итоге к дисфункции органов и тканей. Считается, что основную роль в этих процессах играют миофибробласты, которые дифференцируются под действием профибротических сигналов из фибробластов и других типов клеток стромы.

Анализ имеющихся к настоящему времени данных свидетельствует о том, что мезенхимные стромальные клетки (МСК) принимают участие в регуляции практически всех физиологических процессов, происходящих при репарации и регенерации ткани, в том числе вовлечены в регуляцию фиброза. При этом основная роль в реализации их регуляторных эффектов отводится биологически активным компонентам секрета МСК, включая пул некодирующих регуляторных РНК (нкРНК), которые будучи перенесенными

в составе внеклеточных везикул в другие клетки, способны перепрограммировать их в направлении стимуляции регенеративных процессов. Это делает МСК одним из наиболее перспективных инструментов регенеративной медицины. Однако показано, что в ответ на различные сигналы от поврежденных тканей МСК могут сами дифференцироваться в миофибробласты, а также реагировать специфическими изменениями своего секретома (Grigorieva et al., 2024). При этом остается открытым вопрос о наличии субпопуляций МСК, по-разному отвечающих на профибротические сигналы, и участия в этих процессах нкРНК, секретируемых МСК.

**Цель и задачи работы.** Целью работы является установление механизмов гетерогенности ответа МСК на сигналы профибротического микроокружения с помощью метода scRNA-seq.

Для достижения цели были сформулированы следующие задачи:

1) Выявить субпопуляции МСК, различающиеся по транскриптомным паттернам при культивировании в модели профибротического микроокружения, с помощью метода scRNA-seq.

2) Проанализировать особенности транскриптомного паттерна субпопуляций МСК, которые дифференцируются и не дифференцируются в миофибробласты, идентифицировать в них клеточные типы с использованием ручных и автоматизированных методов типирования и провести анализ регулонов в выявленных кластерах с помощью метода получения регуляторных сетей с целью предположить функциональные характеристики различающихся по транскриптомным паттернам субпопуляций.

3) Построить траектории для определения направления развития МСК, культивируемых в профибротических условиях, с помощью метода упорядочивания единичных клеток вдоль псевдовремени.

4) Установить особенности экспрессии нкРНК в субпопуляциях МСК, по-разному отвечающих на профибротические сигналы.

5) Выявить мембранный белок, экспрессия которого значительно различается в субпопуляциях МСК, по-разному отвечающих на профибротические сигналы, с целью выбора маркера для разделения субпопуляций с использованием клеточного сортирования.

**Научная новизна.** В работе впервые проведен анализ результатов scRNA-seq МСК человека, культивированных в стандартных условиях и в условиях профибротического микроокружения, с использованием современных биоинформатических методов. Предложены новые подходы к идентификации типов клеток в кластерах, находящихся на линии траекторий развития, исходя из знаний о клеточных дифферонах. Впервые было

установлено, что в популяции МСК присутствуют субпопуляции, по-разному отвечающие на профибротические сигналы. Детальная характеристика субпопуляции МСК, которая не дифференцируется в этих условиях в миофибробласты, по дифференциально-экспрессирующимся генам, биологическим процессам и регулонам, позволила сформировать гипотезы о функциональных свойствах этих клеток. Для выявленной субпопуляции установлены характерные поверхностные маркеры, что позволило отсортировать эти клетки для дальнейшего экспериментального изучения и валидации выявленных отличий в их свойствах. Впервые, с помощью метода построения траекторий развития клеточных популяций, для МСК описаны возможные направления их дифференцировки под действием профибротических стимулов, что может быть использовано для контроля дифференцировки МСК в заданном направлении. Весомым дополнением работы стал анализ предшественников микроРНК в выявленных субпопуляциях, что позволило спрогнозировать их участие в регуляции фиброза.

**Теоретическая и практическая значимость исследования.** Применение новых методов биоинформатической обработки результатов анализа scRNA-seq дает уникальные преимущества для изучения гетерогенности популяции клеток, что крайне важно для понимания особенностей ответа клеток на различные стимулы. Результаты диссертационного исследования позволили установить разнообразие ответов МСК в модели, воспроизводящей профибротическое микроокружение. При кластеризации выявлена субпопуляция МСК с потенциально антифибротическими свойствами и проведена ее характеристика. Практической значимостью работы является получение данных, обосновывающих необходимость изучения этой субпопуляции с целью выявления механизмов участия этих клеток в развитии фиброза. Контроль над процессом дифференцировки МСК в фибробласты и миофибробласты, стимуляция регуляторных свойств изучаемой субпопуляции и ее влияния на миофибробласты и фибробласты может позволить предотвращать и обращать развитие фибротических процессов в организме. Новизна и практическая значимость работы подтверждаются полученным патентом РФ №2766707 от 15.03.2022 г. «Средство для лечения фиброза тканей на основе компонентов секретомы мезенхимных стромальных клеток, способ получения и применения средства»

**Методология и методы исследования.** Обработка данных scRNA-seq проводилась с помощью стандартного современного пайплайна, состоящего из этапов контроля качества, картирования, квантификации и нормализации. Последующий анализ полученных первичных данных обрабатывался с помощью пакетов Seurat, SCENIC, scVelo, Velocity, g:Profiler, singleR. В диссертационном исследовании использованы современные методы биоинформатики, клеточной и молекулярной биологии.

### **Основные положения, выносимые на защиту:**

1) Факторы профибротического микроокружения могут определять гетерогенность популяции МСК, реализующейся в формировании клеточных субпопуляций с различными транскриптомными паттернами.

2) С помощью анализа данных РНК секвенирования единичных клеток установлено, что в модели профибротического микроокружения только часть МСК дифференцируется в миофибробласты, остальные клетки обладают транскриптомным профилем, отличным от характерного для миофибробластов.

3) Субпопуляция МСК, не дифференцирующаяся в миофибробласты под действием профибротических стимулов, характеризуется повышенной экспрессией групп генов, отвечающих за организацию и ремоделирование ВКМ, регуляцию метаболических процессов и ангиогенеза, и может быть отделена от других субпопуляций по экспрессии PDGFR $\alpha$ .

4) Гетерогенность ответов МСК на профибротические стимулы может оказывать влияние на развитие фиброза тканей за счет регуляции баланса между субпопуляциями клеток, пополняющих пул миофибробластов, и клеток с предположительно антифибротическими свойствами.

**Степень достоверности результатов проведенных исследований.** Научные результаты диссертационного исследования обладают высокой степенью достоверности, что подтверждается достаточной выборкой, воспроизводимостью результатов измерений, а также согласованностью результатов, полученных с использованием различных методов.

**Апробация работы.** Основные результаты, положения и выводы диссертационного исследования были представлены на 19 научных конференциях и симпозиумах: конгрессе «Single Cells: Technology to Biology» (Сингапур, 24-26 февраля 2019 г.), International Congress Biotechnology: State of the art and perspectives (Москва, 25-27 февраля 2019 г.), Regulatory RNAs (Берлин, 12-14 мая 2019 г.), TERMIS-EU 2019 (Родос, 27-31 мая 2019 г.), IV Национальный конгресс по регенеративной медицине (Москва, 20-23 ноября 2019 г.), VII Молодёжная школа-конференция по молекулярной и клеточной биологии Института цитологии РАН (Санкт-Петербург, 12-15 октября 2020 г.), ISEV 2021 (Франция, 18-21 мая 2021 г.), The 45th FEBS Congress (Словения, 3-8 июля 2021 г.), Noncoding RNA World: From Mechanism to Therapy (Швейцария, 21-23 июля 2021 г.), Moscow Conference on Computational Molecular Biology (MCCMB) (Москва, 30 июля – 2 августа 2021 г.), III Научно-практическая конференция "Секвенирование единичных клеток" (Томск, 23-27 августа 2021 г.), ESGCT Collaborative Virtual Congress 2021 (Бельгия, 19-22 октября 2021 г.), 25-ая Пущинская школа-конференция молодых ученых с международным участием

«Биология – наука XXI века» (Пушино, 18-22 апреля, 2022 г.), IV Научно-практическая школа «Анализ отдельных клеток» (Томск, Томский национальный исследовательский медицинский центр Российской академии наук, 22-26 августа 2022 г.), III Объединенный Научный Форум Физиологов, Биохимиков И Молекулярных Биологов (Сочи, Россия, 3-7 октября 2022 г.), VIII Молодёжная Школа-Конференция по молекулярной биологии и генетическим технологиям Института цитологии РАН (Санкт-Петербург, Россия, 10-14 октября 2022 г.), V Национальный Конгресс по Регенеративной Медицине (Москва, МГУ имени М.В. Ломоносова, 23–25 ноября 2022 г.), 3-я международная конференция "Системная биология и системная физиология: регуляция сложных биологических систем" Внутриклеточная сигнализация и регуляция метаболизма (Москва, Россия, 2-4 декабря 2022), II Международный онлайн-конгресс «Управление старением» (Москва, Россия, 15-16 декабря 2022 г.), European Society of Human Genetics conference (ESHG 2023) (Глазго, United Kingdom, 10-13 июня 2023), Всемирный конгресс: Теория систем, алгебраическая биология, искусственный интеллект: математические основы и приложения (Москва, Россия, 26 июня - 30 августа 2023), 11-я Московская конференция по вычислительной молекулярной биологии (МССМВ) (Москва, Россия, 3-6 августа 2023), МССМВ 2023, Москва, Территория Инновационного Центра “Сколково”, (Россия, 3-6 августа 2023), МССМВ 2023, Москва, Территория Инновационного Центра “Сколково” (Россия, 3-6 августа 2023), Вычислительная биология и искусственный интеллект для персонализированной медицины (Россия, 9-11 августа 2023), Конгресс “CRISPR-2023” (Новосибирск, Россия, 11-13 сентября 2023), TERMIS-AP 2023 (Гонконг, China, 16-19 октября 2023), IV Всероссийская научно-практическая конференция с международным участием «Развитие физико-химической биологии и биотехнологии на современном этапе» (Иркутск, Россия, 25-27 октября 2023), ESHG 2024 - European Human Genetics Conference (Берлин, Germany, 1-4 июня 2024), ESHG 2024 - European Human Genetics Conference (Берлин, Germany, 1-4 июня 2024), Вычислительная биология и искусственный интеллект для персонализированной медицины (Россия, 7-9 августа 2024).

На основании данных диссертации подготовлен и получен патент РФ № 2766707 от 15.03.2022 г.

Материалы диссертационного исследования также апробированы автором в преподавании учебных курсов «Биоинформатика и компьютерные технологии» и «Омиксные технологии» в рамках магистерской программы «Регенеративная биомедицина» по направлению подготовки 06.04.01 «Биология».

**Публикации.** По теме диссертации опубликовано 6 печатных работ в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ по специальности 1.5.8 Математическая биология, биоинформатика.

**Личный вклад автора.** Основные результаты, представленные в диссертационной работе, получены лично автором. Его вклад включает анализ научной литературы, разработку новых вычислительных методов, планирование и проведение вычислительных экспериментов, обработку и интерпретацию данных, подготовку публикаций и участие в научных конференциях. В работе O. Grigorieva et al. 2024 автором была валидирована *in vitro* модель профибротического микроокружения на основе децеллюляризованного внеклеточного матрикса и TGF $\beta$ -1. В исследовании N. Basalova et al. 2023 им был проведен биоинформатический анализ данных секвенирования внеклеточных везикул МСК и выявлены ключевые микроРНК (miR-29c и miR-129), участвующие в регуляции фиброза. В работе A.A. Khozyainova et al. 2023 автор систематизировал методы биоинформатической обработки данных scRNA-seq, включая контроль качества, фильтрацию, кластеризацию и анализ траекторий развития клеточных популяций. В работе O. Grigorieva et al. 2023 автором был выполнен анализ данных scRNA-seq для выявления регуляторных генов дифференцировки МСК в миофибробласты, что позволило установить роль CHD3 и RDH10 в этом процессе. В исследовании M. Arbatsky et al. 2022 им был разработан подход к многомерной визуализации данных методом главных компонент для более полного анализа транскриптомных данных. В работе Basalova et al. 2020 автор провел биоинформатический анализ данных bulkRNA-seq и секвенирования внеклеточных везикул для выявления молекулярных механизмов антифибротического действия МСК.

**Структура и объем работы.** Диссертация состоит из введения, трех глав, заключения, списка литературы, включающего 255 наименований. Диссертация изложена на 223 страницах машинописного текста, содержит 49 рисунков и 26 таблиц.

## **ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

### **Материалы и методы**

МСК подкожной жировой ткани человека были получены от молодых, относительно здоровых доноров и культивированы в стандартных или профибротических условиях, как описано в статьях (Grigorieva O.A. et al., 2023; Grigorieva O.A. et al., 2024). Библиотеки для scRNAseq были подготовлены с помощью 10xGenomics и охарактеризованы, для секвенирования использовали Illumina HiSeq1500 sequencing system (Illumina, Inc., San Diego, CA, USA), как описано в статье (Grigorieva O.A. et al., 2024).

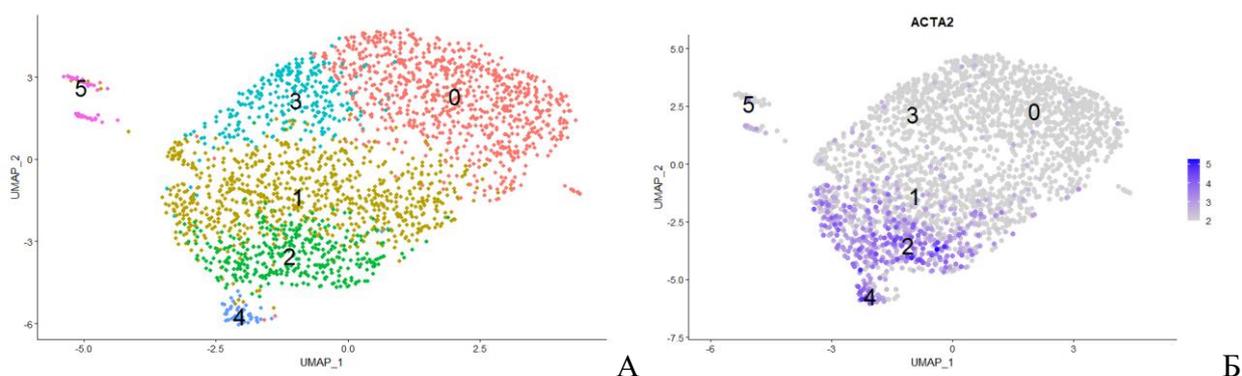
Биоинформатический анализ полученных массивов данных был выполнен с использованием разнообразных программ и пакетов, а также программных сценариев, написанных самостоятельно. Для выравнивания, квантификации, уменьшения размерности, кластеризации, расчета дифференциальной экспрессии, получения файлов barcodes.tsv, features.tsv и matrix.mtx и .cloupe-файлов из первичных данных в формате .fastq был использован cellranger count (v.7). Для получения .aggr-файлов использовали cellranger aggr. Для визуализации и анализа .cloupe-файлов использовали программу Loupe Browser.

Для создания Seurat-объекта из barcodes.tsv, features.tsv и matrix.mtx был использован R-пакет Seurat. Сравнивали результаты таких пайплайнов для анализа данных scRNA-seq, как SCANPY и Asc-Seurat. Интеграцию образцов в режимах CCA, RPCA и SCTransform производили с помощью R-пакета Seurat. Автоматическое типирование клеток проводили с помощью R-пакетов SingleR и celldex. Типирование в ручном режиме осуществляли с использованием клеточных референсов PanglaoDB и CellMarker. Получение .loom-файлов реализовано с помощью пакета velocity, существующего в двух форматах Python и R. Получение траекторий развития клеток образцов осуществляли с помощью Python-пакета scVelo, коллекции R-пакетов Dynverse и веб-платформы Asc-Seurat. Для получения графиков RNA-velocity использовали Python-пакет scVelo. Для работы с данными использовались среды разработки RStudio, Google Colab и PyCharm.

## **Основные результаты и обсуждение**

### **Выявление субпопуляций МСК, по-разному дифференцирующихся под действием профибротических стимулов**

Для моделирования профибротических условий *in vitro* МСК человека культивировали, как описано в статье (Grigorieva O.A. et al., 2024). В результате работы алгоритма анализа результатов scRNAseq получено 6 кластеров в популяции МСК (рисунок 1, А). Было установлено, что в профибротическом микроокружении только часть МСК дифференцируется в миофибробласты, ключевым биомаркером которых является экспрессия  $\alpha$ -гладкомышечного актина ( $\alpha$ -SMA, ген *ACTA2*) (рисунок 1, Б).



**Рисунок 1.** А. Результат кластеризации данных scRNAseq МСК, культивируемых в профибротических условиях (UMAP). Б. Кластер МСК, в которых уровень экспрессии гена *ACTA2* ( $\alpha$ -SMA) в 5 раз больше по сравнению со всеми клетками. N=2.

МСК с высоким уровнем экспрессии *ACTA2* преимущественно располагаются во 2 и 4 кластерах (рисунок 1, Б), которые далее именуется  $\alpha$ SMA<sup>+</sup>-кластерами. Кластеры 0, 1, 3 и 5, в которых сосредоточились МСК с низким уровнем экспрессии *ACTA2* далее именуется  $\alpha$ SMA<sup>-</sup>-кластерами.

Для функционального анализа 2 (таблица 1) и 4 (таблица 2) кластеров были использованы гены с  $\text{avg\_log}_2\text{FC} > 0.6$  (31 ген). В этих кластерах заметно выражен ответ на повреждение и процессы регуляции заживления ран и клеточной миграции, что является основными процессами при развитии фиброза. Также обнаруживаются гены, участвующие в мышечном сокращении.

**Таблица 1.** Функциональное обогащение белков 2 кластера  $\alpha$ SMA<sup>+</sup>-субпопуляции МСК.

GO термин	Описание	Количество в сети	Сила	Доля ложных отклонений
GO:0006936	Muscle contraction	7/248	1.31	7.16e-05
GO:0009611	Response to wounding	6/532	0.98	0.0397

Примечание: **GO-term** (генная онтология), **description** (название термина), **count in network** (первая цифра – количество генов в анализируемой сети, совпавших со списком всего термина), **strength** ( $\text{Log}_{10}(\text{наблюдаемый}/\text{ожидаемый})$ ). Мера величины эффекта обогащения. Это соотношение между количеством белков в вашей сети, аннотированных термином, и количеством белков, которые, могли бы быть аннотированы этим термином в случайной сети того же размера. **false discovery rate** (значительность обогащения, р-значения, скорректированные на множественное тестирование в каждой категории с помощью процедуры Бенджамина-Хохберга)

**Таблица 2.** Функциональное обогащение белков 4 кластера  $\alpha$ SMA<sup>+</sup>-субпопуляции МСК.

GO термин	Описание	Количество в сети	Сила	Доля ложных отклонений
GO:0061041	Regulation of wound healing	4/140	1.27	0.0454
GO:0030334	Regulation of cell migration	6/865	0.68	0.0454

Для проверки наличия дифференцированных миофибробластов было изучено распределение экспрессии известных маркеров этих клеток в субпопуляциях МСК ( $\alpha$ -SMA, Desmin, Vimentin, CRBP1, Calponin, Paladin 41 g, Podoplanin, Stromelysin-3,

Endosialin, Gamma-SMA, P4H (prolyl-4 hydroxylase), Cadherin-11, GB-42, Tropomyosin-1, Thyl-1 и Cofilin) (De Wever et al., 2008; Pho et al., 2008). Перечисленные маркеры преимущественно сосредоточены во 2 и 4 кластерах массива, что подтверждает дифференцировку МСК этих кластеров в миофибробласты.

**Особенности транскриптомного паттерна субпопуляции МСК, которая не дифференцируется в миофибробласты в профибротических условиях.**

Для того, чтобы предположить отличающиеся функциональные свойства субпопуляции МСК, которая при культивировании в профибротическом микроокружении не демонстрировала транскриптомный паттерн, характерный для миофибробластов, были изучены особенности экспрессии генов в 0, 1, 3 и 5 кластерах. Для получения интерактома 0 кластера были использованы гены с  $avg\_log2FC > 0.4$  (31 ген). Важным процессом в данном кластере является GO:0030198 Extracellular matrix organization,  $3.87e-05$  (8 генов из 338 генов этого класса GO). Для функционального анализа 1 кластера было использовано 52 гена (таблица 3). Важными процессами в кластере являются организация ВКМ, ангиогенез, оссификация и развитие соединительной ткани, что подтверждает имеющиеся данные о направлениях дифференцировки МСК в фибробласты, гладкомышечные клетки, адипоциты, хондробласты и остеобласты. Участие МСК в регуляции ангиогенеза, по-видимому, может быть опосредовано действием секрета этих клеток и их способностью стабилизировать образующиеся сосуды.

**Таблица 3.** Наиболее представленные процессы в клетках 1 кластера aSMA-субпопуляции МСК в терминах GO.

GO термин	Описание	Количество в сети	Сила	Доля ложных отклонений
GO:0030198	Extracellular matrix organization	22/338	1.39	7.00E-21
GO:0072359	Circulatory system development	18/872	0.89	1.43E-08
GO:0001568	Blood vessel development	13/500	0.99	9.23E-07
GO:0001501	Skeletal system development	12/499	0.96	5.88E-06
GO:0009888	Tissue development	20/1760	0.63	7.57E-06
GO:0001503	Ossification	9/265	1.11	2.21E-05
GO:0048514	Blood vessel morphogenesis	10/410	0.96	6.13E-05
GO:0061448	Connective tissue development	8/223	1.13	7.24E-05

Для функционального анализа 3 кластера было использовано 24 гена, 5 кластера – 20 (таблица 4). В 5 кластере активно происходит пролиферация клеток (GO:0051301 Cell division  $7.58e-17$ ).

**Таблица 4.** Наиболее представленные процессы в клетках 3 кластера  $\alpha$ SMA<sup>-</sup> субпопуляции МСК в терминах GO.

GO термин	Описание	Количество в сети	Сила	Доля ложных отклонений
GO:0006695	Cholesterol biosynthetic process	6/41	2.08	4.67E-08
GO:0046890	Regulation of lipid biosynthetic process	7/197	1.46	1.75E-06
GO:0062012	Regulation of small molecule metabolic process	8/449	1.16	6.41E-06
GO:0050810	Regulation of steroid biosynthetic process	5/89	1.66	2.70E-05
GO:0045338	Farnesyl diphosphate metabolic process	3/5	2.69	4.27E-05
GO:0045540	Regulation of cholesterol biosynthetic process	4/43	1.88	9.41E-05

Для того, чтобы охарактеризовать выделенные субпопуляции в них были определены дифференциально экспрессирующиеся гены. В списке высокоэкспрессированных генов  $\alpha$ -SMA<sup>-</sup> субпопуляции представлено 20 генов (таблица 5).

**Таблица 5.** Сравнение кратности изменения дифференциально представленных генов в  $\alpha$ SMA<sup>-</sup>-субпопуляции МСК со всеми клетками образца (0, 1, 3 и 5 кластеры).

Ген	avg log2FC	Кластер	Ген	avg log2FC	Кластер
CTSK	1.550	0	PDGFRA	0.445	0
CTSL	0.551/0.276	0/5	SOX4	0.427	0
MMP1	0.269	0	SPON2	0.397	0
MMP3	0.827	0	PLPP3	0.393	0
FBLN1	0.627	0	DPP4	0.387	0
MARCKSL1	0.618	0	HCFC1R1	0.372	0
LINC01705	0.585	0	TSHZ2	0.343	0
CHI3L1	0.541	0	COLEC12	0.316	0
GPNMB	0.539	0	PLTP	0.291	0
VIM	0.470/0.269	0/3	CLEC2B	0.357	0

Гены CTSK (Buhling et al., 2004) и SH3PXD2B (de Vos et al., 2020) участвуют в разрушении компонентов ВКМ. DPT (Unamuno et al., 2020) обеспечивает связь между поверхностью фибробластов кожи и их внеклеточным окружением. SNX9 (Wilkes et al., 2015) участвует в разрушении EGFR-рецептора. KLHL24 (Alkhalifah et al., 2018) необходим для поддержания баланса между стабильностью и разрушением промежуточных филаментов, обеспечивающих целостность поверхности кожи. FBLN1 (Liu et al., 2019) играет роль в клеточной адгезии и транспорте по волокнам ВКМ. Интересно, что среди высокоэкспрессированных генов в этой субпопуляции присутствует ген MMP11 (stromelysin-3), участвующий в активации Notch-сигнального пути (Mukhi S.,

2011), активация которого способствует миофибробластной дифференцировке (Aoyagi-Ikeda et al., 2011). Возможно, такой эффект можно объяснить началом незавершенной дифференцировки МСК в миофибробласты. Высокоэкспрессирующиеся гены входят в состав кластеров GO\_BP, объединяющих гены, отвечающие за организацию ВКМ, что характеризует  $\alpha$ -SMA<sup>-</sup>-субпопуляцию МСК как сдерживающую фибротические изменения путем ремоделирования образующихся при развитии фиброза белков ВКМ (таблица 6).

**Таблица 6.** Кластеризация по GO\_BP (Gene ontology, biological process) высокоэкспрессированных генов (> 2 раз) в  $\alpha$ -SMA<sup>-</sup>-субпопуляции МСК. T – размер кластера GO\_BP. Q – количество анализируемых генов. T/Q – число генов из анализируемых, входящих в данный кластер.

GO термин	Описание	Количество в сети	Padj
GO:0030154	Cell differentiation	32/4440	2.927*10 <sup>-2</sup>
GO:0045597	Positive regulation of cell differentiation	13/894	3.008*10 <sup>-2</sup>
GO:0097435	Supramolecular fiber organization	12/739	2.261*10 <sup>-2</sup>
GO:0030198	Extracellular matrix organization	9/399	2.522*10 <sup>-2</sup>
GO:0043062	Extracellular structure organization	9/400	2.572*10 <sup>-2</sup>

Дополнительно был проведен анализ высокоэкспрессированных  $\alpha$ -SMA<sup>+</sup>-субпопуляции МСК генов (таблицы 7, 8).

**Таблица 7.** Сравнение кратности изменения дифференциально представленных генов в  $\alpha$ -SMA<sup>+</sup>-субпопуляции МСК со всеми клетками образца (2 и 4 кластеры).

Ген	avg log2FC	Кластер	Ген	avg log2FC	Кластер
IGFBP3	1.953	2	CSRP1	1.215	4
CNN1	1.145	2	TM4SF1	0.942	4
SERPINE1	0.960	2	SYNPO2	0.902	4
MYLK	0.947/1.470	2/4	EDIL3	0.899	4
KRT7	0.911	2	LMOD1	0.860	4
CLIC3	0.909	2	CD36	0.828	4
PPME1	0.743	2	TPM2	1.062/ 0.823	2/4
ACTG2	2.103	4	MYH11	0.816	4
APOE	1.921	4	DSTN	0.792	2
COL4A1	1.339	4	CRYAB	0.779	4
PPP1R14A	1.218	4			

MYLK (Russo et al., 2005), CNN1 (Borkham-Kamphorst et al., 2014), MYL9 (Iwasaki et al., 2001), PPP1R14A (является ингибитором PPP1CA) участвуют в сокращении гладкомышечных клеток. SYNPO2 (Kai et al., 2015) участвует в формировании сократительных стрессовых волокон актина, ориентированных параллельно длинной оси клетки. LMOD1 (Woczowska et al., 2015), CDC42EP3 (Hirsch et al., 2001) участвуют в организации актиновых филаментов. CSRP2 заметно снижается в ответ на PDGF-BB или клеточное повреждение, что запускает деление гладкомышечных клеток и их

дифференцировку. Высокоэкспрессирующиеся гены входят в состав кластеров GO\_BP, объединяющих гены, отвечающие за регуляцию организации актиновых волокон, что характерно для направления дифференцировки МСК в миофибробласты (таблица 8).

**Таблица 8.** Кластеризация по GO\_BP (Gene ontology, biological process) высокоэкспрессированных генов (> 2 раз) в  $\alpha$ -SMA<sup>+</sup>- субпопуляции МСК. T – размер кластера GO\_BP. Q – количество анализируемых генов. T/Q – число генов из анализируемых, входящих в данный кластер.

GO термин	Описание	Количество в сети	Padj
GO:0030029	Actin filament-based process	18/824	$1.473 \cdot 10^{-4}$
GO:0030036	Actin cytoskeleton organization	17/723	$1.258 \cdot 10^{-4}$
GO:0097435	Supramolecular fiber organization	15/739	$5.248 \cdot 10^{-3}$
GO:0007015	Actin filament organization	13/449	$5.250 \cdot 10^{-4}$
GO:0006936	Muscle contraction	10/368	$2.465 \cdot 10^{-2}$
GO:0110053	Regulation of actin filament organization	9/288	$2.296 \cdot 10^{-2}$
GO:1902905	Positive regulation of supramolecular fiber organization	8/216	$2.140 \cdot 10^{-2}$
GO:0051495	Positive regulation of cytoskeleton organization	8/234	$3.821 \cdot 10^{-2}$
GO:0030239	Myofibril assembly	5/68	$4.142 \cdot 10^{-2}$

#### **Идентификация клеточных типов в выявленных кластерах МСК с использованием ручных и автоматизированных методов типирования**

**Автоматическое типирование.** Возможности автоматического типирования клеток на сегодняшний день крайне ограничены. В широко используемом пакете SingleR есть возможность использования нескольких клеточных референсов. При использовании референса HumanPrimaryCellAtlasData в популяции МСК, культивированных в стандартных условиях, определено 7 типов клеток (преобладающий тип - фибробласты, а также индуцированные стволовые клетки, макрофаги, мезенхимальные стволовые клетки, нейроны, гладкомышечные клетки и тканевые стволовые клетки), при использовании BlueprintEncodeData – 6 типов (адипоциты, фибробласты, макрофаги, мезангиальные клетки, миоциты и гладкомышечные клетки). При использовании референса HumanPrimaryCellAtlasData в популяции МСК, культивируемых в профибротических условиях, определено 8 типов клеток (хондроциты, эндотелиальные клетки, фибробласты, индуцированные стволовые клетки, мезенхимальные стволовые клетки, остеобласты, гладкомышечные клетки и тканеспецифичные стволовые клетки), при использовании BlueprintEncodeData – 5 типов (хондроциты, эндотелиальные клетки, фибробласты, миоциты и гладкомышечные клетки). Полученные данные соответствуют тканевому источнику МСК, используемых в работе (подкожная жировая ткань). Стоит отметить, что в профибротических условиях не выявлены в популяции МСК индуцированные стволовые

клетки и макрофаги, определяемые в контрольном образце, а также детектированы хондроциты, что, возможно, говорит о том, что под влиянием профибротических сигналов часть МСК может дифференцироваться в хондроциты. Однако, автоматическое типирование рекомендовано использовать для ознакомления с массивом данных, и его применение обосновано в основном для клеток крови или нервной ткани. Для типирования клеток соединительной ткани предпочтительнее использовать общепринятые специфические маркеры.

Типирование клеток в ручном режиме проводилось по высокоэкспрессирующимся в кластере специфическим маркерам отдельных типов клеток. В программе Loupe Browser, являющейся приложением программного конвейера Cell Ranger, можно экспортировать таблицы, в которых указана сравнительная экспрессия генов клеток данного кластера по отношению к клеткам всего образца. В R-пакете Seurat также есть возможность получения списка дифференциально экспрессирующихся генов с помощью функции **FindMarkers**. По умолчанию Seurat выполняет дифференциальное выражение на основе непараметрического критерия суммы рангов Уилкоксона, выводя список наиболее дифференциально экспрессированных генов. Используя данную функцию, был получен список дифференциально экспрессированных генов для всех кластеров популяции МСК, культивированных в стандартных условиях, и проведено сравнение отобранных алгоритмом генов с клеточными референсами Cell Marker и PanglaoDB. Из 6 кластеров контрольного образца по специфическим маркерам удалось идентифицировать только два кластера как МСК. При типировании клеток кластеров МСК, культивируемых в профибротических условиях, идентифицируется большее число типов клеток, но это не дает целостного и полного понимания того, как клетки реагируют на стимулы микроокружения, какие процессы протекают в клетках в данный момент и какова траектория развития этих клеток.

Для решения этого вопроса необходимо рассматривать не только дифференциально экспрессированные гены, специфические для определенных типов клеток, но и выраженные биологические процессы на момент секвенирования. Поскольку большинство клеток в процессе дифференцировки находятся в промежуточных состояниях, их затруднительно типировать по специфическим маркерам. Для типирования клеток по биологическим процессам мы использовали базы данных STRING (Szklarczyk et al., 2019). В базе данных есть возможность построения интерактома наиболее представленных генов в кластере. В режиме *multiple proteins* импортируется список высокопредставленных генов. Полученные результаты позволили получить дополнительную информацию о не типированных кластерах (таблицы 9, 10).

**Таблица 9.** Кластеризация по биологическим процессам МСК, культивируемых в стандартных условиях.

Кластер	GO:BP	Гены
0	GO:0006412 Translation	RPS12, RPS28, TMA7, THBS1, RPL35A, RPS27L, RPL12, RPLP2
1	Fibroblasts	MMP3, SFRP4, PI16
2	Smooth muscle cells	TIMP1, ACTA2, TPM2
3	GO:0000278 Mitotic cell cycle	TOP2A, STMN1, MKI67, CENPF, SMC4, UBE2C, TUBB4B, UBE2S, TYMS, RRM2, CKS2, TPX2, BIRC5, TUBA1B
4	GO:0006936 Muscle contraction	TPM2, ACTG2, MYL9, CRYAB, MYL12A, GSTO1
5	GO:0030198 Extracellular matrix organization	TIMP1, POSTN, LOX, SPARC, VCAN, COL3A1, MFAP5, FBN1, LUM, FN1, DCN, CLU
6	GO:0006955 Immune response	CHI3L1, CCL2, CTSD, FUCA1, CTSL, CTSK, HLA-DRA, ANXA1, CXCL8, CLU, PSAP, MDK, CTSS, CTSS

**Таблица 10.** Кластеризация по биологическим процессам МСК, культивируемых в профибротических условиях.

Кластер	GO:BP	Гены
0	GO:0030334 Regulation of cell migration	SOD2, MDK, CXCL12, DCN, FBLN1, SERPINF1, FGF7, IL1R1, GPNMB
	GO:0030198 Extracellular matrix organization	COL6A3, FTH1
	GO:1901342 Regulation of vasculature development	CHI3L1
1	GO:0030198 Extracellular matrix organization	SERPINE1, PTX3, THBS1, SPARC, LOX, FN1, COL1A1, MFAP5, COL11A1, TGFBI, SERPINH1, LOXL2
	GO:0001568 Blood vessel development	THY1, SERPINE1, THBS1, SPARC, LOX, FN1, COL1A1, COL11A1, TGFBI, LOXL2
	GO:0048514 Blood vessel morphogenesis	
	GO:0072359 Circulatory system development	
	GO:0001503 Ossification	COL1A1, LOX, SPARC, VCAN, COL1A2, TPM4, COL11A1, COL5A2, IGFBP3
GO:0001501 Skeletal system development	MYH9, COL1A1, LOX, THBS1, VCAN, CDH2, COL1A2, TPM4, FN1, TPM1, COL5A1, CAP1, ITGBL1, LOXL2, MYO1B, PALLD	
2	GO:0006936 Muscle contraction	MYL12A, MYL9, TPM1, MYLK, CALD1, TPM2, ACTA2, MYL6
	GO:0010757 Negative regulation of plasminogen activation	SERPINE1, THBS1
3	GO:0006695 Cholesterol	SQLE, HMGCS1, INSIG1, FDPS, ACAT2,

	biosynthetic process	FDFT1
4	GO:0006936 Muscle contraction	MYL9, TPM1, MYLK, CALD1, LMOD1, TPM2, MYH11, ACTG2, ACTA2, CRYAB, MYL6
	GO:0042060 Wound healing	MYH9, HSPB1, MYL9, ACTB, TPM1, CSRP1, CD151, CD36
	GO:0061041 Regulation of wound healing	MYH9, APOE, CAV1, MYLK, CD36
5	GO:0051301 Cell division	CCNB1, TPX2, BIRC5, TUBA1B, TUBB, UBE2C, SMC4, CENPF, CDC20, PTTG1, PRC1, STMN1

### **Анализ выявленных регулонов в кластерах МСК с помощью метода получения регуляторных сетей**

Регулонами называется группа генов, регулируемых и коэкспрессирующихся с определенным транскрипционным фактором (ТФ). Из ТФ регулонов 0 кластера МСК, культивируемых в профибротических условиях, ELK3 является негативным регулятором транскрипции, PRRX2 играет роль в безрубцовом заживлении ран, RUNX2 участвует в остеобластной дифференцировке и формировании скелета, NFATC4 принимает участие в адипогенной дифференцировке и формировании мышечных волокон. Из регулонов 1 кластера ARID5B участвует в адипогенезе, ATF5 в выживании клеток, их пролиферации и дифференцировке. Из регулонов 2 кластера среди генов регулона 2 кластера отсутствуют гены, участвующие в регуляции процесса фиброза. Из регулонов 3 кластера TBX5 контролирует количество предшественников мезенхимальных клеток и хондроцитов. Из регулонов 4 кластера SEBPA участвует в дифференцировке адипоцитов. Из регулонов 5 кластера E2F1 останавливает дифференцировку адипоцитов (рисунок 2).



**Рисунок 2.** Выявленные регулоны в кластерах МСК, культивируемых в профибротических условиях.

Для того, чтобы понять, в каких процессах участвуют гены, регулируемые различными ТФ, были получены списки этих генов для каждого ТФ. После этого для каждой группы генов была выполнена кластеризация по GO. Для того чтобы предположить отличия в функциональных свойствах  $\alpha$ SMA<sup>+</sup> субпопуляции МСК, были выбраны кластеры 1 и 3 для дальнейшего анализа. В первом кластере карты белок-белковых взаимодействий были построены для ТФ TEAD3 и ARID5B, так как количество регулируемых ими генов достаточно для построения графа. Для ТФ TEAD3 были отобраны процессы, которые потенциально могут влиять на патогенез фиброза, включая сборку актиновых нитей и организацию цитоскелета, дифференцировку клеток, развитие эпителия и межклеточные взаимодействия (таблица 11). Для ТФ ARID5B было отобрано три процесса: отрицательная регуляция ответа на внешние стимулы, дифференцировки и биосинтетических процессов, что может указывать на угнетение биосинтетических процессов в МСК в ответ на профибротические стимулы (таблица 12).

**Таблица 11.** Биологические процессы для генов, регулируемых ТФ TEAD3.

GO термин	Описание	Количество в сети	Сила	Доля ложных отклонений
GO:0030029	Actin filament-based process	10 of 592	0.78	0.0093
GO:0030036	Actin cytoskeleton organization	9 of 516	0.79	0.0128
GO:0000094	Cell morphogenesis involved in differentiation	9 of 566	0.75	0.0193
GO:0060429	Epithelium development	12 of 1109	0.59	0.0294
GO:0098609	Cell-cell adhesion	8 of 505	0.75	0.0422

**Таблица 12.** Биологические процессы для генов, регулируемых ТФ ARID5B.

GO термин	Описание	Количество в сети	Сила	Доля ложных отклонений
GO:0048585	Negative regulation of response to stimulus	37 of 1636	0.27	0.0143
GO:0045596	Negative regulation of cell differentiation	21 of 728	0.38	0.0184
GO:0031327	Negative regulation of cellular biosynthetic process	35 of 1592	0.26	0.0312

В третьем кластере карты белок-белковых взаимодействий были построены для ТФ TBX15 и MSC. Среди генов, контролируемых ТФ TBX15, два гена входят в сигнальный путь, регулирующий актиновый цитоскелет, что так же может объяснять разницу в уровне экспрессии  $\alpha$ SMA между субпопуляциями МСК (таблица 13). Для ТФ MSC особый интерес представляет сигнальный путь, регулирующий ответ на внешние стимулы, который, по-видимому, вовлечен в адаптацию клеток третьего кластера к изменениям микроокружения. Активность данного пути, наряду с выраженным биосинтезом холестерина, может указывать на специализированную функцию этих клеток в поддержании мембранной стабильности и их способности эффективно реагировать на внешние сигналы (таблица 14).

**Таблица 13.** Сигнальный путь для генов, регулируемых ТФ TBX15.

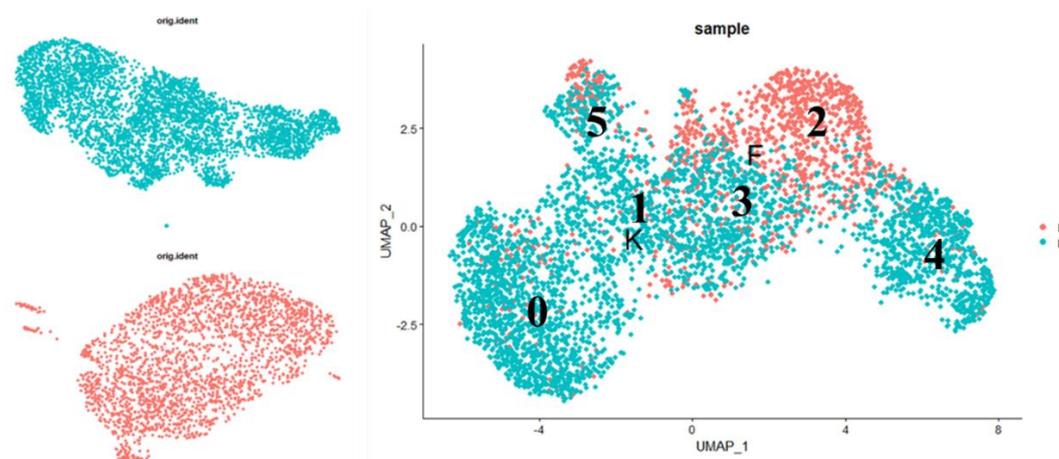
Сигнальный путь	Описание	Количество в сети	Сила	Доля ложных отклонений
hsa04810	Regulation of actin cytoskeleton	2 of 209	1.97	0.0389

**Таблица 14.** Сигнальные пути для генов, регулируемых ТФ MSC.

Сигнальный путь	Описание	Количество в сети	Сила	Доля ложных отклонений
HSA-9648895	Response of EIF2AK1 (HRI) to heme deficiency	3 of 15	2.21	0.0028
HSA-96333012	Response of EIF2AK4 (GCN2) to amino acid deficiency	4 of 99	1.52	0.0077
HAS-8953897	Cellular response to external stimuli	6 of 558	0.94	0.0347

### **Построение траектории для определения направления развития МСК, культивируемых в профибротических условиях, с помощью метода упорядочивания единичных клеток вдоль псевдовремени**

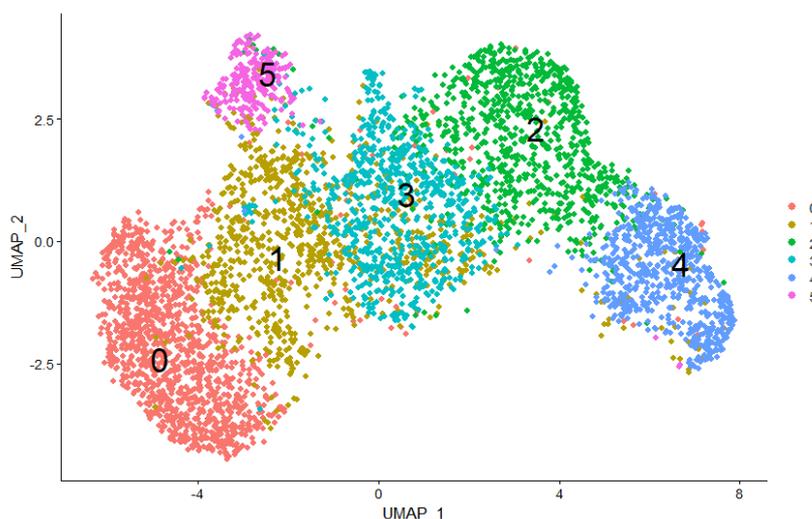
Для анализа направлений развития и дифференцировки клеток недостаточно анализа отдельных образцов, так как необходимо провести сравнение клеток нескольких образцов друг с другом для изучения не пересекающихся кластеров клеток (рисунок 3).



**Рисунок 3.** Создание интегрированного объекта для анализа транскриптома единичных МСК, культивируемых в стандартных (К) или профибротических (F) условиях.

При визуализации интегрированного объекта было обнаружено, что МСК, культивируемые в стандартных и профибротических условиях, перекрываются только в 1 и 3 кластерах. 2 и 5 кластеры представляют особый интерес, так как это кластеры клеток, которых не было в объектах отдельных образцов (рисунок 4).

Изучение траекторий развития клеток позволяет не только анализировать дифференциальную экспрессию, но также дает возможность проследить направление развития или дифференцировки клеток в ответ на сигналы от микроокружения (рисунок 5).

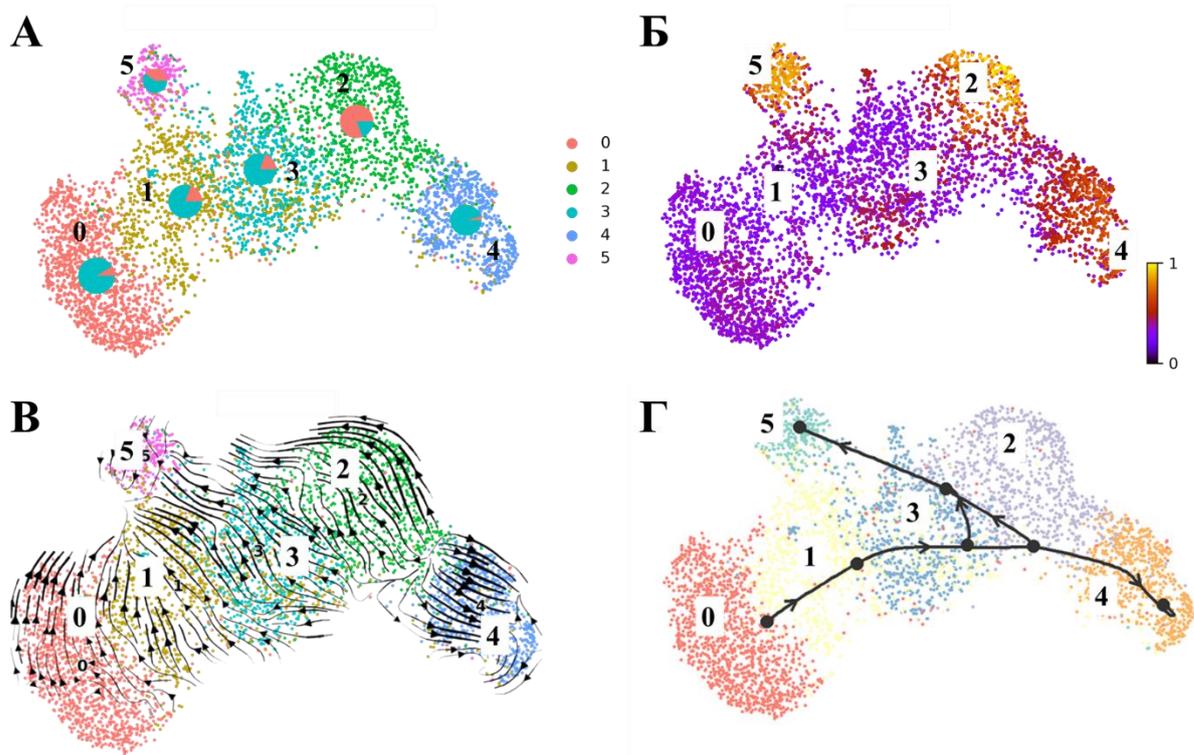


**Рисунок 4.** UMAP-график кластеризации интегрированного объекта для анализа транскриптома единичных МСК, культивируемых в стандартных или профибротических условиях. 0 – Тип клеток не определен, характерных процессов нет; 1 – Миофибробласты, HSA-445355 Smooth Muscle Contraction; 2 – Фибробласты, GO:0042127 Regulation of cell population proliferation, GO:0005201 Extracellular matrix structural constituent; 3 – Базальные клетки, фибробласты, HSA-445355 Smooth Muscle Contraction; 4 – активно делящиеся клетки, GO:0051301 Cell division, GO:0000278 Mitotic cell cycle; 5 – Базальные клетки, фибробласты, GO:0030199 Collagen fibril organization, GO:0030198 Extracellular matrix organization, GO:0061448 Connective tissue development, GO:0042060 Wound healing, GO:1905049 Negative regulation of metalloproteinase activity.

Так, одна из предполагаемых траекторий в интегрированном образце (рисунок 5, Г) начинается с клеток, представленных в 0 кластере, где четкое определение клеточного типа отсутствует, что говорит о наличии стволовых или низкодифференцированных клеток, готовых к дальнейшему развитию. Затем траектория развития направляется к кластеру 1, где присутствуют миофибробласты, отвечающие за сокращение гладкой мускулатуры (HSA-445355). Эти клетки обладают высокой пластичностью и могут трансформироваться в другие типы клеток в ответ на сигналы окружающей среды, что связано с процессами регенерации и заживления. После этого траектория проходит через кластер 3 с фибробластами и базальными клетками, которые активно участвуют в регуляции клеточной пролиферации и формировании ВКМ (GO:0042127, GO:0005201). Наконец, траектория завершается в 4 кластере, состоящем из активно делящихся клеток, задействованных в клеточном делении (GO:0051301) и митотическом цикле (GO:0000278). Это конечное состояние указывает на активный рост и восстановление ткани, обеспечивая жизнеспособность клеточной популяции.

Вторая траектория, начинающаяся во 2 кластере с фибробластами, также проходит через базальные клетки в 3 кластере и завершает путь в 5 кластере с фибробластами и базальными клетками. Эта траектория акцентирует внимание на процессах организации коллагена (GO:0030199), развития соединительной ткани (GO:0061448), заживления ран (GO:0042060) и негативной регуляции активности металлопротеиназ (GO:1905049). Тканевая регенерация и восстановление зависят от организованной работы внеклеточной матрицы, что предоставляет структурную поддержку и способствует восстановлению поврежденных участков.

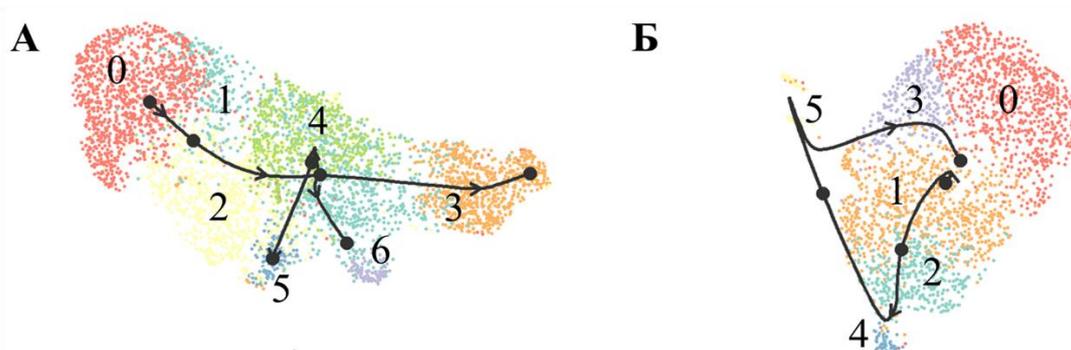
График RNA-velocity (рисунок 5, В) повторяет основные направления траекторий с уточнениями направлений внутри кластеров. Дополнительные графики с указанием соотношения количества клеток разных образцов в каждом кластере и график латентного времени показывают, что наибольшее изменение пропорций клеток происходит в кластерах 2 и 5 (рисунок 5, А), которые на графике латентного времени имеют наиболее желтый оттенок (рисунок 5, Б), что говорит о том, что в данных кластерах находятся наиболее дифференцированные формы клеток.



**Рисунок 5.** Анализ клеток интегрированного объекта. **А.** Результат кластеризации клеток интегрированного объекта с t-SNE, обозначающими соотношение клеток контрольного и профибротического образцов в каждом кластере (контроль – голубой, фиброз - коралловый); **Б.** График латентного времени клеток интегрированного объекта (положение клеток на временной линии дифференцировки или развития); **В.** Векторы RNA-velocity клеток интегрированного объекта, показывающие направления дифференцировки клеток внутри кластеров; **Г.** Paga plot с траекториями развития клеток интегрированного образца, показывающими направления дифференцировки клеток между макросостояниями (кластерами).

Методы построения траекторий развития позволяют улучшить результаты типирования клеток, основываясь на информации о дифферонах и положении нетипированного кластера на траектории развития между типированными кластерами, а также литературных данных, описывающих морфологические свойства и транскриптом начальных, промежуточных и конечных форм клеток, связанных преемственной линией дифференцировки. При анализе траекторий развития МСК, культивируемых в стандартных условиях, не было обнаружено нетипированных кластеров, находящихся в промежуточном положении между кластерами с известными типами клеток (рисунок 6 А). Анализ траекторий развития МСК, культивируемых в профибротических условиях, показал, что в 0 кластере располагаются клетки, похожие по транскриптому на гладкомышечные клетки (ГМК), характеризующиеся активацией процессов организации ВКМ (GO:0030198) и регуляцией развития сосудов (GO:1901342). 1 кластер также связан с процессами организации ВКМ (GO:0030198) и морфогенезом сосудов (GO:0048514), что указывает на активное участие клеток этих кластеров в ангиогенезе. Вероятно, эти клетки

представляют собой предшественников ГМК, участвующих в формировании сосудистой сети. Клетки 2 кластера активны в таких процессах, как сокращение мышц (GO:0006936) и негативная регуляция активации плазминогена (GO:0010757), что может свидетельствовать о механизмах, отвечающих за сосудистый тонус и предотвращение тромбообразования. Клетки 3 кластера, активированные в процессе биосинтеза холестерина (GO:0006695), вероятно, играют важную роль в липидном обмене и поддержании клеточной мембраны, что может быть связано с регуляцией клеточной функциональности и выживанием. 4 кластер, состоящий из клеток, ответственных за сокращение мышц (GO:0006936) и заживление ран (GO:0042060), указывает на активацию клеток в ответ на повреждающие стимулы. Эти клетки обладают высокой пластичностью и могут участвовать в восстановительных процессах. 5 кластер, ассоциированный с делением клеток (GO:0051301), демонстрирует активные процессы пролиферации, которые необходимы для замещения поврежденных клеток и обеспечения регенерации тканей (рисунок 6, Б).



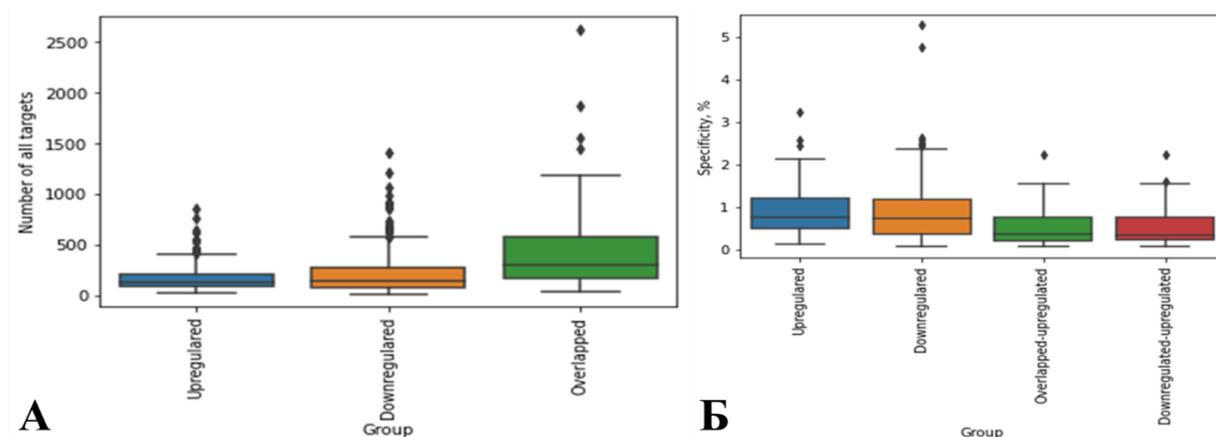
**Рисунок 6.** А. Траектория развития МСК, культивируемых в стандартных условиях. 0 - GO:0006412 Translation; 1 - Fibroblasts (MMP3); 2 - Smooth muscle cells (TIMP1, ACTA2, TPM2); 3 - GO:0000278 Mitotic cell cycle; 4 - Fibroblasts (TPM2); 5 - GO:0030198 Extracellular matrix organization; 6 - GO:0006955 Immune response; Б. Траектория развития МСК, культивируемых в профибротических условиях. 0 - Smooth muscle cells (ACTA2) GO:0030198 Extracellular matrix organization & GO:1901342 Regulation of vasculature development; 1 - GO:0030198 Extracellular matrix organization & GO:0048514 Blood vessel morphogenesis; 2 - GO:0006936 Muscle contraction & GO:0010757 Negative regulation of plasminogen activation; 3 - GO:0006695 Cholesterol biosynthetic process; 4 - GO:0006936 Muscle contraction & GO:0042060 Wound healing; 5 - GO:0051301 Cell division

### **Установление особенностей экспрессии нкРНК в субпопуляциях МСК, поразному отвечающих на профибротические сигналы**

В результате анализа генов в единичных МСК в соответствии с принадлежностью к кластерам выявленных терминов и путей GO, DO, KEGG и Reactome, связанных с фибротическими состояниями, было отобрано 16 генов со сниженной и 9 генов с

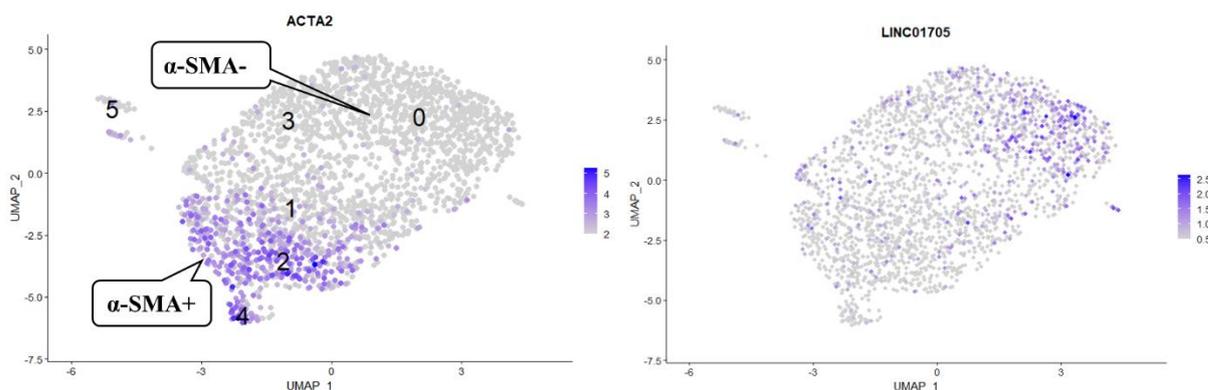
повышенной в профибротических условиях экспрессией. Затем был осуществлён поиск микроРНК, для которых отобранные гены являлись мишенями, и создано два списка, в один из которых входили микроРНК, имеющие в качестве мишеней гены с повышенной экспрессией, в другой - гены с пониженной экспрессией. Затем по полученным спискам была произведена обратная процедура поиска всех возможных генов-мишеней данных микроРНК. Были найдены все микроРНК, для которых данные гены являются мишенями, и поделены на 3 группы: downregulated - 254 микроРНК против группы генов с экспрессией сниженной в профибротических условиях, upregulated - 127 микроРНК против группы генов с повышенной экспрессией, overlapped - 51 микроРНК, имеющие мишенями гены из обеих групп. Были подсчитаны общее количество мишеней каждой микроРНК, а также их специфичность по отношению к отобранным генам. Медианные значения специфичности составили 0,8% для групп upregulated и downregulated, и 0,4% в группе overlapped (рисунок 7, Б). Медианные значения по числу мишеней составляют 150 мишеней для групп downregulated и upregulated, и примерно 300 мишеней для группы overlapped (рисунок 7, А).

Для сужения круга микроРНК был проведен семантический анализ и обогащения их генов-мишеней с фильтрацией по терминам, связанным с процессами регуляции процесса фиброза и МСК ('stem', 'fibrosis', 'wound', 'differentiation', 'actin', 'matrix', 'fibroblast', 'myofibroblast', 'aging', 'senescence', 'apoptosis', 'wounding', 'muscle', 'programmed', 'extracellular', 'encapsulating', 'proliferation', 'mesenchymal'). Дополнительно был проведен анализ литературы для формирования окончательного списка в кандидатных микроРНК для изучения их влияния на субпопуляции МСК: hsa-mir-10a-5p, hsa-mir-34a-5p, hsa-mir-27a-3p, hsa-mir-194-5p, hsa-mir-18a-5p, hsa-mir-20a-5p, hsa-mir-451a, hsa-mir-129-5p, hsa-mir-29a-3p, hsa-mir-29b-3p, hsa-mir-29c-3p.



**Рисунок 7.** А. Бокс-плот, отображающий количество мишеней микроРНК; Б. Бокс-плот, отображающий процент специфичности микроРНК к отобранным генам-мишеням.

Среди дифференциально представленных некодирующих РНК в субпопуляции МСК, не дифференцирующихся в миофибробласты в условиях профибротического микроокружения, была выявлена LINC01705 (рисунок 8). По данным ENSEMBLE (Yates et al., 2020) данная последовательность относится к длинным некодирующим РНК (днРНК), имеет 3 транскрипта: LINC01705-201 (562 пн.), LINC01705-202 (1070 пн.), LINC01705-203 (863 пн.). Интересно, что по данным базы данных LncExpDB (Li et al., 2021) эта РНК чаще всего встречается в МСК легких и яичек.



**Рисунок 8.** Представленность днРНК LINC01705 в МСК, культивируемых в профибротических условиях.

Анализ экспрессии с помощью базы данных Expression Atlas показал, что LINC01705 имеет низкую экспрессию (не выше 3 TPM) в 14 здоровых тканях, наибольшую экспрессию в choroid plexus (сосудистое сплетение желудочков мозга), при этом она высоко представлена (более 12 TPM) в поврежденных тканях при раке поджелудочной железы и некоторых других онкологических заболеваниях.

Ab initio предсказание взаимодействия с белками для каждого транскрипта показало, что LINC01705-201 взаимодействует с 73 белками. Наилучшие показатели  $p$ -value  $< 0.001$  продемонстрировали белки FLG, SRRM2 и ZNF469. LINC01705-202 - 96 белками. Наилучшие показатели  $p$ -value  $< 0.001$  продемонстрировали белки ZNF469, PCF11 и NCL. LINC01705-203 - 716 белками. Наилучшие показатели  $p$ -value  $< 0.001$  продемонстрировали белки ZNF469, FLG и SPEN. Таким образом, белок ZNF469 (zinc finger protein 469) проявляет лучшее взаимодействие со всеми транскриптами. По данным NCBI данный белок отвечает за синтез волокон коллагена.

Для оценки возможности участия LINC01705 в регуляции процессов, связанных с фиброзом, был получен список микроРНК, с которыми эта днРНК может взаимодействовать. После этого среди генов-мишеней каждой из отобранных 13 микроРНК было получено количество генов, относящихся к одному из терминов GO:

extracellular structure organization, response to wounding, tissue morphogenesis, tissue remodeling, wound healing. У hsa-mir-335-5p больше всего мишеней в 3 из 5 перечисленных процессов, а процессы tissue remodeling, wound healing управляются наибольшим количеством микроРНК. Среди генов, участвующих в регуляции фиброза и таргетируемых наибольшим количеством микроРНК - VEGFA, RTN2, POLM, GATAD1, SPRED1. Некоторые из 13 микроРНК связаны с такими заболеваниями, как цирроз печени и фиброз желчного пузыря. Интересно, что по данным базы полногеномного поиска ассоциаций (GWAS) (Riat et al., 2019) LINC01705 ассоциируется с 23 вариантами, которым соответствуют 9 заболеваниям, включая связанные с фиброзом, например, развитие келоидных рубцов. Недавно была выявлена функция LINC01705 как эндогенной конкурирующей miR-186-5p, регулирующей экспрессию белка TRP при раке груди (Du et al., 2020). В одной из работ приведен список возможных основных регуляторов эпителиально-мезенхимального перехода и TGF $\beta$  сигнального пути при протоковой аденокарциноме поджелудочной железы: TGFB2-AS1, AL138930.1, LINC01705, AC245041.1, UCA1, и NKILA (Lina et al., 2020). Предполагают участие lncRNA-LINC01705-201 (ENSG00000232679.2, ENST00000438158.1) в эпителиально-мезенхимальном переходе пигментного эпителия сетчатки (Yang et al., 2021).

#### **Определение мембранных белков с целью выбора маркера для разделения субпопуляций МСК, по-разному отвечающих на профибротические сигналы, с использованием клеточного сортирования**

Для разделения клеточных субпопуляций с дальнейшим анализом их функциональных свойств с помощью клеточного сортирования необходимо выявить мембранные маркеры этих субпопуляций. Из базы данных UniProt были экспортированы таблицы, содержащие информацию о локализации белков в клетке. Сопоставив список белков с этой базой, было определено 13 белков, являющиеся мембранными. Для сортирования клеток  $\alpha$ -SMA<sup>-</sup> - субпопуляции может быть предложен белок PDGFR $\alpha$ , так как он преимущественно распределен именно в данной субпопуляции. Для сортирования клеток  $\alpha$ -SMA<sup>+</sup> - субпопуляции МСК из 10 найденных белков были предложены LIMS2, CRIM1 и CDH2, так как они преимущественно распределены именно в данной субпопуляции.

### **ЗАКЛЮЧЕНИЕ**

Анализ транскриптома единичных МСК человека, культивируемых в стандартных или профибротических условиях, был проведен с использованием платформы 10x Genomics. Применение набора инструментов для анализа scRNA-seq (контроль качества, нормализация, шкалирование, выбор высоковариабельных генов, снижение размерности,

кластеризация, визуализация, дифференциальная экспрессия, типирование, RNA-velocity, построение траекторий развития, анализ регулонов, анализ регуляции с помощью нкРНК) позволило установить гетерогенность ответа МСК на сигналы профибротического микроокружения. Так, была выявлена субпопуляция МСК, которая не дифференцируется в этих условиях в миофибробласты. На основании полученных данных можно предположить, что эта субпопуляция обладает уникальными функциональными свойствами, которые могут оказывать влияние на процесс развития фиброза. Анализ регуляции с помощью нкРНК выявил в этой субпопуляции специфические миРНК и днРНК, обладающих потенциальным влиянием на процесс развития фиброза. Установленные характерные поверхностные маркеры для данной субпопуляции МСК открывают перспективы для дальнейшего изучения её биологических свойств идентификации потенциальных биомаркеров с целью диагностики заболеваний и перспективных молекулярных мишеней для терапии фибротических заболеваний.

#### **Основные результаты и выводы**

1) С помощью метода scRNA-seq выявлены субпопуляции МСК, дифференцирующиеся и не дифференцирующиеся в миофибробласты при культивировании в модели профибротического микроокружения.

2) Анализ транскриптомного паттерна субпопуляций МСК позволил типировать в выявленных кластерах миофибробласты, хондроциты, эндотелиальные клетки, фибробласты, миоциты и гладкомышечные клетки, а также предположить, что субпопуляция МСК, которая не дифференцируется в миофибробласты в профибротических условиях, преимущественно вовлечена в процессы, связанные с организацией и ремоделированием ВКМ, регуляцией ангиогенеза, обмена липидов, дифференцировкой клеток в остеогенном и хондрогенном направлениях. При этом по результатам анализа регулонов в клетках этой субпопуляции была установлена активность ряда транскрипционных факторов, участвующих в подавлении фибротического ответа в ответ на повреждение, стимуляции жизнеспособности, пролиферации и дифференцировки клеток.

3) Анализ траекторий развития МСК, культивируемых в профибротических условиях, подтвердил гетерогенность ответа клеток на сигналы от микроокружения в виде разделения на различные по транскриптомному паттерну клеточные субпопуляции стромальных клеток.

4) Анализ экспрессии нкРНК в субпопуляциях МСК, по-разному отвечающих на профибротические сигналы, позволил выявить днРНК (LINC01705) и ряд микроРНК (hsa-mir-10a-5p, hsa-mir-34a-5p, hsa-mir-27a-3p, hsa-mir-194-5p, hsa-mir-18a-5p, hsa-mir-20a-5p,

hsa-mir-451a, hsa-mir-129-5p, hsa-mir-29a-3p, hsa-mir-29b-3p, hsa-mir-29c-3p), потенциально вовлеченные в регуляцию дифференцировки и функциональных свойств МСК, не дифференцирующихся в миофибробласты.

5) Были установлены мембранные белки (PDGFR $\alpha$ , LIMS2, CRIM1 и CDH2), экспрессия которых значимо различается в субпопуляциях МСК, по-разному отвечающих на профибротические сигналы, что позволяет предложить их в качестве маркеров для разделения субпопуляций с использованием клеточного сортирования.

## СПИСОК ПУБЛИКАЦИЙ

**Список публикаций по теме диссертации в рецензируемых изданиях, опубликованных в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ по специальности 1.5.8 Математическая биология, биоинформатика.**

1. Grigorieva O., Basalova N., Dyachkova U., Novoseletskaya E., Vigovskii M., **Arbatskiy M.**, Kulebyakina M., Efimenko A. Modeling the profibrotic microenvironment in vitro: model validation // **Biochemical and Biophysical Research Communications**. – **2024**. Vol. 733. P. 150574. EDN: HQSAIE, Импакт-фактор 2,2 (JIF). (0,84/0,15)<sup>1</sup>.
2. Basalova N., **Arbatskiy M.**, Popov V., Grigorieva O., Vigovskiy M., Zaytsev I., Novoseletskaya E., Sagaradze G., Danilova N., Malkov P., Cherniaev A., Samsonova M., Karagyaur M., Tolstoluzhinskaya A., Dyachkova U., Akopyan Z., Tkachuk V., Kalinina N., Efimenko A. Mesenchymal stromal cells facilitate resolution of pulmonary fibrosis by miR-29c and miR-129 intercellular transfer // **Experimental and Molecular Medicine**. – **2023**. – Vol. 55. – № 7. – pp. 1399–1412. EDN: KSCMMI. Импакт-фактор 12,9 (JIF). (0,78/0,1).
3. Khozyainova A.A., Valyaeva A.A., Arbatsky M.S., Isaev S.V., Iamshchikov P.S., Volchkov E.V., Sabirov M.S., Zainullina V.R., Chechekhin V.I., Vorobev R.S., Menyailo M.E., Tyurin-Kuzmin P.A., Denisov E.V. Complex Analysis of Single-Cell RNA Sequencing Data // **Biochemistry (Moscow)** – **2023**. – Vol. 88. – № 2. – pp. 231–252. EDN: QFSJMW. Импакт-фактор 2,2 (JIF). (1,26/0,3).
4. Grigorieva O., Basalova N., Vigovskiy M., **Arbatskiy M.**, Dyachkova U., Kulebyakina M., Kulebyakin K., Tyurin-Kuzmin P., Kalinina N., Efimenko A. Novel Potential Markers of Myofibroblast Differentiation Revealed by Single-Cell RNA Sequencing Analysis of Mesenchymal Stromal Cells in Profibrotic and Adipogenic Conditions // **Biomedicines**. –

---

<sup>1</sup> В скобках приведён объём публикации в печатных листах и вклад автора в печатных листах

2023. – Vol. 11. – № 3. – P. 840. doi: 10.3390/biomedicines11030840. EDN: OUZBFK.  
Импакт-фактор 3,9 (JIF). (0,96/0,15).

5. **Arbatsky M.**, Tyurin-Kuzmin P., Kulebyakin K., Chechekhin V., Kalinina N., Sysoeva V., Semina E., Rubina K. Points of Significance: Principal Component Analysis for Biocentric Data Visualization // **BioNanoScience**. – 2022. – Vol. 12. – pp. 1366–1380. doi: 10.1007/s12668-022-01021-w. EDN: QAKCPE. Импакт-фактор 3,2 (JIF). (0,84/0,2).

6. Basalova N., Sagaradze G., **Arbatskiy M.**, Evtushenko E., Kulebyakin K., Grigorieva O., Akopyan Z., Kalinina N., Efimenko A. Secretome of Mesenchymal Stromal Cells Prevents Myofibroblasts Differentiation by Transferring Fibrosis-Associated microRNAs within Extracellular Vesicles // **Cells**. – 2020. – Vol. 9. – №5. – P. 1272, EDN: ESMODC. Импакт-фактор 5,2 (JIF). (0,84/0,2).