



ФЕДЕРАЛЬНОЕ
ГОСУДАРСТВЕННОЕ
УЧРЕЖДЕНИЕ «ФЕДЕРАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
«ФУНДАМЕНТАЛЬНЫЕ
ОСНОВЫ БИОТЕХНОЛОГИИ»
РОССИЙСКОЙ АКАДЕМИИ НАУК»

119071, Москва, Ленинский пр-т, д. 33, стр. 2
Тел. +7 (495) 954-52-83, факс (495) 954-27-32
www.fbras.ru, info@fbras.ru

07.09.2023 № 25-04-19/685
На № 11615-6210/2023 от 24.07.2023

УТВЕРЖДАЮ
ИО Директора ФИЦ
биотехнологии РАН,
д.б.н. Федоров А.Н.



О Т З Ы В

**ведущей организации на диссертационную работу Менделевич Аси Владимировны
"Статистические вопросы, связанные с техническими и биологическими
вариациями, возникающие при аллель-специфическом анализе данных
секвенирования", представленную на соискание ученой степени кандидата
физико-математических наук по специальности
1.5.8. – Математическая биология, биоинформатика.**

Актуальность. В данных секвенирования присутствуют множественные источники вариации, включая биологические, экспериментальные и вычислительные. Все они вносят свой вклад в общий шум, который накапливается в процессе выборки, обработки данных и множества других неучтенных факторов. Понимание источников шума критически важно для корректной интерпретации полученных результатов.

Вопреки значимости этого вопроса, многие из существующих инструментов для анализа аллель-специфической экспрессии существенно недооценивают технический шум, что систематически приводит к ложным открытиям. Самым распространённой моделью в исследованиях аллельного дисбаланса является биномиальная. Эта модель совсем не учитывает наличие дополнительной вариации. Таким образом, разработка методов, учитывающих техническую избыточную дисперсию, представляет собой важную и актуальную научную проблему.

В данной работе представлены методы, позволяющие вычислять техническую избыточную дисперсию, происходящую из экспериментального и вычислительного процессирования данных, что позволяет делать точные оценки аллельного дисбаланса и корректным образом проводить дифференциальный анализ аллель-специфической экспрессии. Более того, протоколы, предложенные в последней главе, не требуют существенных дополнительных затрат при проведении эксперимента. Разработанные программные продукты имеются в свободном доступе на github. В свете вышеизложенного, диссертационная работа Менделевич А.В. является крайне актуальной.

Структура работы. Изложение следует классической схеме, текст диссертации состоит из введения и четырёх глав: обзора литературы, и трёх глав, содержащих описание методов и результатов. Введение содержит постановку задачи и описывает актуальность и новизну работы. Обзор литературы содержит историю развития области исследования аллельного дисбаланса, также в нём рассматриваются существующие методы измерения уровня экспрессии и аллельного дисбаланса, и обсуждаются методы оценки шума в экспериментах РНК-секвенирования. В главе 2 обсуждается вычислительный подход для учета технического шума при количественной оценке аллельного дисбаланса, основанный на обработке данных технической репликации. В главе 3 с его помощью изучается полногеномное воздействие ингибитора метилтрансферазы 5-аза-2'-деоксицитидина на различные клональные линии, и показывается, что метилирование ДНК является одним из ключевых механизмов поддержания моноаллельной аутосомной экспрессии. В главе 4 приводится экспериментальный протокол и усовершенствованные вычислительные инструменты, позволяющие масштабировать анализ данных аллель-специфической экспрессии без существенных затрат. Диссертацию завершают заключение, выводы и список литературы.

Научная новизна. В ходе исследования было установлено, что существующие инструменты недостаточно эффективны для точной количественной оценки аллель-специфической экспрессии. Это повлекло за собой необходимость разработки новых подходов, позволяющих проводить дифференциальный анализ аллель-специфической экспрессии. Было показано, что техническая компонента избыточной дисперсии неотделима от биологической вариабельности без технической репликации, и также было выявлено, что количество реплик менее важно, чем метод обработки данных из этих реплик. Помимо вычислительного протокола, был предложен новый экспериментальный дизайн, который обеспечивает возможность проведения точного количественного анализа данных аллель-специфической экспрессии экономным и масштабируемым способом. Наконец, по результатам исследования полногеномных эффектов применения ингибитора метилтрансферазы 5-аза-2'-деоксицитидина на различных клеточных линиях было показано, что метилирование ДНК играет ключевую роль в регуляции ASE.

Научная значимость результатов и их практическое значение. Диссертация затрагивает актуальные вопросы разработки протокола для точного дифференциального анализа аллель-специфической экспрессии. Результаты исследования подчеркивают важность корректного учета технической избыточной дисперсии для повышения воспроизводимости результатов при работе с аллель-разрешенными данными РНК-секвенирования и избежания завышенного уровня ложноположительных результатов. Следование предложенным экспериментальным и вычислительным протоколам для обработки данных позволяет достигать большей статистической корректности, без существенного увеличения затрат на эксперимент. Этот подход может быть полезен во многих транскриптомных исследованиях и быть обобщён на другие типы данных.

Замечания к работе. В работе не было исследовано влияние протокола секвенирования РНК на поведение предложенных моделей (например, широко распространенный вариант захвата РНК за кэп), что несколько ограничивает применимость результатов. Также доклад давал довольно неполное представление от

диссертационной работе в части положений, выносимых на защиту, что было рекомендовано исправить.

Заключение. Диссертационная работа Менделевич Аси Владимировны «Статистические вопросы, связанные с техническими и биологическими вариациями, возникающие при аллель-специфическом анализе данных секвенирования» является законченным научно-квалификационным исследованием, в котором решена задача разработки методов точной количественной оценки дифференциальной аллель-специфической экспрессии, а также адаптации протокола для экономного и масштабируемого анализа данных ASE. Решение этой задачи имеет большое значение для повышения воспроизводимости результатов в исследованиях аллельного дисбаланса. По совокупности полученных результатов работа Менделевич Аси Владимировны соответствует п.7 «Положения о порядке присуждения ученых степеней» ВАК МОН Российской Федерации, утвержденных Постановлением Правительства РФ № 74 от 30 января 2002 года в редакции от 20.06.2011 №475. Результаты достоверны и опубликованы в рецензируемых журналах, индексируемых базами WoS и Scopus (Q1), автореферат адекватно отражает содержание диссертации. Автор заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 1.5.8. — математическая биология, биоинформатика.

Материалы диссертации были рассмотрены и обсуждены на Межлабораторном семинаре ФИЦ биотехнологии РАН 26 июня 2023 года (протокол № 1).

кандидат биологических наук,
старший научный сотрудник,
руководитель группы
регуляторной транскриптомики и эпигеномики
Федерального Исследовательского Центра
«Фундаментальные основы биотехнологии»
Российской Академии Наук

/Медведева Ю. А./

12 июля 2023 года

Подпись старшего научного сотрудника, руководителя группы регуляторной транскриптомики и эпигеномики, кандидата биологических наук Медведевой Юлии Анатольевны удостоверяю.

Ученый секретарь ФИЦ биотехнологии РАН
к.б.н. Орловский А.Ф.

