

**ОТЗЫВ официального оппонента
на диссертацию на соискание ученой степени
кандидата биологических наук Рябых Григория Кирилловича
на тему: «РНК-хроматиновые взаимодействия: базы данных,
интегративный анализ и функциональная аннотация»
по специальности 1.5.8. Математическая биология, биоинформатика**

Общая характеристика работы и актуальность исследования

Представленная к защите диссертация Г.К. Рябых посвящена актуальной проблеме современной биологии изучения РНК-хроматиновых взаимодействий. Эта область характеризуется противоречием между высокой биологической значимостью некодирующих РНК (нкРНК), быстрым накоплением экспериментальных данных и недостатком унифицированных инструментов для их систематизации, сравнения и ограниченность интегративных подходов, связывающих карты РНК-хроматиновых контактов с эпигенетическими метками, архитектурой хроматина и профилями экспрессии.

Для изучения РНК-хроматиновых взаимодействий разработаны высокопроизводительные методы, позволяющие картировать контакты как конкретной РНК (ОТА), так и всех РНК (АТА) с локусами хроматина. Эти подходы сформировали большой массив публичных данных, но поскольку соответствующие экспериментальные данные были обработаны при помощи различных биоинформатических сценариев анализа данных, то сопоставление полученных результатов затруднено.

В диссертационной работе автором решена комплексная задача по разработке методологических основ, инструментов и аналитической инфраструктуры для систематического изучения РНК-хроматинового интерактома.

Научная новизна и практическая значимость работы

Научная новизна диссертационного исследования не вызывает сомнений. В диссертационной работе впервые разработан и применен

стандартизированный вычислительный протокол для обработки исходных экспериментальных данных всех основных методов изучения РНК-хроматинового интерактома, что обеспечило сопоставимость ранее разрозненных наборов данных.

Создана первая специализированная аналитическая база данных RNA-Chrom. База данных поддерживается веб-интерфейсом для интерактивного анализа и визуализации. Практическая востребованность RNA-Chrom подтверждается тем, что ресурс уже использовался в независимых исследованиях для решения задач более высокого уровня, включая анализ cis-регуляторных длинных нкРНК и построение интегративных моделей РНК–белок–ДНК взаимодействий.

Предложена и апробирована интегративная схема, связывающая физические РНК-хроматиновые контакты (RNA-Chrom) с данными о корреляциях между экспрессией нкРНК и эпигенетическими метками (HiMoRNA). На примере конкретных нкРНК (PVT1, MIR31HG и др.) показано, что такая интеграция позволяет формулировать интерпретируемые гипотезы о механизмах эпигенетической регуляции экспрессии генов длинными некодирующими РНК.

Впервые проведен полномасштабный сравнительный анализ существующих данных РНК-хроматинового интерактома, позволивший количественно охарактеризовать и сопоставить различные экспериментальные методы по воспроизводимости и специфичности результатов.

Степень обоснованности положений, выносимых на защиту, научных выводов и рекомендаций

Положения, выносимые на защиту хорошо обоснованы и подтверждены результатами биоинформатического анализа. Автор применил современные и адекватные поставленным задачам методы исследования и стек программных технологий.

Хотелось бы отметить высокое качество разработанных автором базы RNA-Chrom и ее веб-интерфейса.

На основе полученных результатов автором предложена стратегия анализа данных РНК-хроматинового интерактома, которая включает ценные методические рекомендации для анализа таких данных.

Достоверность результатов

Все экспериментальные данные, использовавшиеся в работе, находятся в открытом доступе, и результаты их анализа могут быть воспроизведены.

Результаты работы по теме диссертации опубликовано 6 печатных работ, из них 5 статей в рецензируемых научных изданиях.

Оформление диссертации и автореферата

Диссертация оформлена в соответствии с требованиями, предъявляемыми к работам такого рода. Структура диссертации логична и включает введение, обзор литературы, описание материалов и методов, изложение результатов и их обсуждение, заключение, основные результаты и выводы, список литературы и большое количество приложений.

Текст написан грамотным научным языком, изложение материала последовательное и логичное. Список литературы полон и включает все основные современные работы по теме исследования, что свидетельствует о хорошем знании автором состояния проблемы.

Содержание автореферата соответствует содержанию диссертации.

Соответствие диссертации паспорту специальности

Диссертационная работа Г.К. Рябых полностью соответствует паспорту научной специальности 1.5.8 – Математическая биология, биоинформатика (биологические науки). Представленные в диссертации результаты принадлежат направлениям исследования:

2. Компьютерная системная биология (геномика, транскриптомика,

протеомика, метаболомика, другие омиксные исследования).

11. Организация, ведение и использование специализированных мультидисциплинарных банков данных и баз знаний по биологии и медицине, в т.ч. банков междисциплинарных данных.

12. Разработка и применение новых вычислительных алгоритмов для анализа экспериментальных данных в биологии и медицине.

Замечания к диссертационной работе

Есть небольшие замечания по оформлению диссертационной работы:

1. Нет заключения по обзору литературы.
2. Формулы не пронумерованы.
3. Изредка используются жаргонизмы, например, стр. 40, Рисунок 5. "доля побед".
4. Задачи опосредованным образом соответствуют выводам: задач 7, а выводов 5. Тем не менее в выводах содержится решение всех поставленных задач.

Также следует отметить некоторые аспекты, требующие дополнительного обсуждения или уточнения:

1. Раздел "2.1.2.8. Аннотация РНК-частей контактов генами".

Согласно данным CAGE-seq и FANTOM5 кроме генов, РНК транскрибируется с множества других районов. Имело бы смысл, если РНК, образующая контакт не соответствуют аннотированным генам, проводить ее дополнительную аннотацию по сайтам инициации транскрипции из FANTOM5.

2. стр. 48

Учитывая значительную перепредставленность ближних контактов (РД-скейлинг), в дальнейших анализах в данной статье не включались

взаимодействия, находящиеся ближе 1 Мб от генов, кодирующих соответствующие РНК.

Не теряется ли таким образом важная информация?

Это артефакт метода или это имеет какой-то биологический смысл и специально отобрано в ходе эволюции?

4. Раздел "3.1.2. Разработка БД и структура данных"

- лучше было поместить этот раздел в "Материалы и методы";
- не описаны поля таблиц в базе данных. В частности, не понятно, как устроена структура UpstreamDownstreamGenes в таблице Contacts. Таблица Conatacts - является ключевой, но ее поля и как она используется - не описано.

5. Раздел "3.1.3. Разработка веб-сервиса RNA-Chrom" также лучше было поместить в "Материалы и методы".

6. стр. 62

С биологической точки зрения RNA-Chrom переводит анализ РНК-хроматиновых взаимодействий из режима описания отдельных контактов в режим приоритизации биологически осмысленных гипотез.

Известны ли успешные такие примеры?

7. Логичнее было бы поместить раздел 3.1.5 перед разделом 3.1.4.

8. стр. 80

При расширении контактов на +/- 1, +/- 5, +/- 10, +/- 25 и +/- 50 тысяч п.н. процент РНК, у которых пики из HiMoRNA подтверждаются хотя бы одним контактом, увеличивается до 38.5, 42.7, 45.7, 50.1 и 53% соответственно.

В разделе "3.4.2. Сравнение реплик в данных АТА" автор указывает, что 5000 п.н. являются эмпирической оценкой точности позиционирования контакта в методах АТА.

Понятно, что при расширении районов, даже чисто по случайным причинам, будет увеличиваться количество подтвержденных пиков из HiMoRNA.

Какие биологические механизмы могут обеспечивать такие дальние взаимодействия?

Какова вероятность подтверждения пика при таком количестве РНК-хроматиновых взаимодействий при расширении районов до 25 и 50 тыс. п.н.?

Какой процент пиков должен подтверждаться исходя из биологических предположений и текущих наших знаний об этих процессах? Ведь не только нкРНК участвуют в регуляции гистоновых меток. Традиционно считается, что основную роль в этом играют транскрипционные факторы.

9. стр. 82

Помимо этого, мы выявили функциональную связь этих днРНК с другими эпигенетическими метками: NR2F1-AS1 – H3K4me2, MIR31HG – H3K79me2, SNHG1 – H3K4me3, PVT1 – H3K4me1, H3K4me3, H3K9me3 (рис. 27А)

Как подтверждается, что это именно функциональная связь?

10. Раздел "2.4.4. Воспроизводимость (конкордантность) контактов в репликах". Почему не использовался стандартный подход для воспроизводимости в репликах на основе IDR (Irreproducible Discovery Rate)?

11. Отдельным вопросом для обсуждения является функциональность и биологическая значимость выявляемых РНК-хроматиновых взаимодействий. Безусловно, есть очень биологически значимые РНК-хроматиновые взаимодействия, например, XIST.

Однако,

- мРНК и нкРНК всего ~300 000 молекул на всю клетку, включая ту, которая в цитоплазме;
- соответственно молекул нкРНК, с учетом их, как правило низкой экспрессии, всего 1-2 на всю клетку, т.е. будет 1-2 контакта;
- в рекомендациях автор предлагает отбирать РНК с числом контактов >10 000.

Из этого следует:

- почему нужно **ОЧЕНЬ** большое количество клеток и прочтений, чтобы была какая-то воспроизводимость выявляемых локусов РНК-хроматиновых взаимодействий;
- с биологической точки зрения >99% выявляемых РНК-хроматиновых контактов не должно существенно влиять на регуляцию хроматина и экспрессии. Действительно, если в клетке 1-2 транскрипта соответствующей нкРНК и >10 000 то вероятность того, что нужная РНК свяжется с нужным контактом $< 1-2/10\ 000$.
- биологический смысл могут нести преимущественно ближние взаимодействия, где выше вероятность что соответствующая нкРНК свяжется с нужным локусом.

Заключение

Вместе с тем, указанные замечания не умаляют значимости диссертационного исследования. Диссертация отвечает требованиям, установленным Московским государственным университетом имени М.В.Ломоносова к работам подобного рода. Содержание диссертации соответствует специальности 1.5.8. Математическая биология, биоинформатика (по биологическим наукам), а также критериям, определенным пп. 2.1-2.5 Положения о присуждении ученых степеней в Московском государственном университете имени М.В.Ломоносова. Диссертационное исследование оформлено согласно требованиям Положения о совете по защите диссертаций на соискание ученой степени кандидата наук,

на соискание ученой степени доктора наук Московского государственного университета имени М.В.Ломоносова.

Таким образом, соискатель Рябых Григорий Кириллович заслуживает присуждения ученой степени кандидата биологических наук по специальности 1.5.8. Математическая биология, биоинформатика.

Официальный оппонент:

Доктор биологических наук,

Научный руководитель,

направление «Вычислительная биология»,

Научный центр генетики и наук о жизни,

Научно-технологический университет «Сириус»

Колпаков Федор Анатольевич

05.06.2026

Федор Анатольевич Колпаков
подпись

И.С.

Контактные данные:

тел.: [REDACTED], e-mail: [REDACTED]

05.06.2026

Специальность, по которой официальным оппонентом защищена диссертация:

1.5.8 Математическая биология, биоинформатика

Адрес места работы:

354340, Краснодарский край, федеральная территория «Сириус», ул.

Олимпийский проспект, д. 1,

Автономная некоммерческая образовательная организация высшего

образования «Научно-технологический университет «Сириус», научный

центр генетики и наук о жизни, направление «Вычислительная биология»

Тел.: [REDACTED]