

ОТЗЫВ

официального оппонента на диссертацию

Летягиной Анны Евгеньевны «**Анализ влияния последовательности, расположенной после сайта полиаденилирования, на уровень зрелой мРНК репортёрного гена eGFP в культивируемых клетках человека HEK293T**», представленную к защите на соискание ученой степени кандидата биологических наук по специальности 1.5.7 – «Генетика».

Актуальность темы исследования

Тема диссертационного исследования А.Е. Летягиной, посвященная анализу влияния последовательности, расположенной после сайта полиаденилирования (DSE), на уровень зрелой мРНК репортёрного гена eGFP в культивируемых клетках человека HEK293T, является высоко актуальной и соответствует современным направлениям молекулярной биологии и генетики. Изучение регуляторных элементов, расположенных в 3'-областях генов, остаётся менее разработанным по сравнению с промоторными и энхансерными регионами, несмотря на их важную роль в процессинге пре-мРНК и терминации транскрипции. Работа вносит существенный вклад в понимание механизмов, посредством которых нуклеотидный состав и структурные особенности DSE модулируют эффективность экспрессии генов у млекопитающих. Полученные результаты имеют не только фундаментальное значение для расширения знаний о посттранскрипционной регуляции, но и практическую ценность для биотехнологии и геной инженерии, открывая возможности для рационального дизайна терминаторов транскрипции с заданными свойствами. Использование современных методов, таких как массовый параллельный репортёрный анализ (МПРА) и машинное обучение, подчёркивает новаторский характер исследования и его соответствие современным научным трендам.

Научная новизна и значимость работы

В диссертационной работе впервые на основе метода массового параллельного репортёрного анализа (МПРА) системно исследовано влияние нуклеотидной последовательности элемента, расположенного ниже сигнала полиаденилирования (DSE), и стабильности его предсказанной вторичной структуры на уровень зрелой мРНК репортёрного гена eGFP в культивируемых клетках человека HEK293T. Автором разработана и применена оригинальная вычислительная модель DoSIA (Downstream Sequence element Iterative Analyzer), способная предсказывать уровень экспрессии гена на основе последовательности DSE, что позволило рационально сконструировать и экспериментально подтвердить функциональность новых синтетических терминаторов транскрипции. Полученные результаты вносят существенный вклад в фундаментальные представления о регуляции терминации транскрипции РНК-полимеразы II у млекопитающих и открывают перспективы для целенаправленного дизайна регуляторных элементов с заданной эффективностью, что имеет важное практическое значение для геной инженерии и биотехнологии, в частности для оптимизации экспрессии целевых генов в клеточных системах.

Структура и краткое содержание работы

Диссертация представляет собой системное исследование влияния последовательности Downstream Sequence Element (DSE), расположенной после сигнала полиаденилирования (СПА), на уровень зрелой мРНК репортёрного гена eGFP в

культивируемых клетках человека НЕК293Т. Методологическая стратегия основана на применении массового параллельного репортёрного анализа (МПРА), который позволил оценить влияние тысяч переменных последовательностей DSE на экспрессию гена. В ходе работы проведена оптимизация условий ПЦР для минимизации образования химерных молекул, разработан и применён программный инструмент MPRAdecoder для обработки данных высокопроизводительного секвенирования. Установлено, что наибольшее влияние на уровень зрелой мРНК оказывает область +17...+40 п.н. после СПА, а низкая стабильность предсказанной вторичной структуры РНК в этой области коррелирует с высокой экспрессией репортёра. На основе данных МПРА обучена модель машинного обучения DoSIA, способная предсказывать уровень экспрессии на основе последовательности DSE, и использована для рационального дизайна синтетических терминаторов, модулирующих экспрессию гена. Результаты работы расширяют представления о регуляторной роли DSE в терминации транскрипции и предлагают практические инструменты для конструирования генетических конструкций с заданным уровнем экспрессии.

Работа изложена на 208 страницах, иллюстрирована 29 рисунками и 10 таблицами. Библиография включает 535 источников. Структура включает: введение, литературный обзор, описание материалов и методов, результаты, заключение, выводы, благодарности, список литературы и приложения.

Оформление диссертации соответствует установленным правилам научного цитирования, библиографические ссылки оформлены корректно. Содержание работы соответствует пунктам 4 и 11 Паспорта научной специальности 1.5.7. – «генетика» (биологические науки). Содержание диссертации полностью отражено в автореферате.

Обоснованность и достоверность результатов

Научные положения и выводы диссертационной работы характеризуются высокой степенью обоснованности. Методологическая основа исследования сочетает современные экспериментальные подходы массового параллельного репортёрного анализа (МПРА) с биоинформатическими методами машинного обучения и моделированием вторичных структур РНК. Экспериментальная часть включает оптимизированные протоколы ПЦР (в том числе эмульсионной ПЦР для минимизации химерных продуктов), временную трансфекцию культивируемых клеток НЕК293Т, выделение тотальной РНК, обратную транскрипцию, количественную ПЦР, флуоресцентную микроскопию. Для анализа данных МПРА использован самостоятельно разработанный программный инструмент MPRAdecoder, а также обученные модели машинного обучения, что позволило количественно оценить влияние нуклеотидных последовательностей после сигнала полиаденилирования на уровень зрелой мРНК. Статистическая обработка данных, кросс-валидация моделей, воспроизводимость результатов в независимых экспериментах и использование контрольных конструкций подтверждают надёжность и достоверность полученных результатов.

Результаты исследования представлены в 3 публикациях в рецензируемых научных журналах, входящих в международные базы цитирования, а также доложены на 9 научных конференциях.

Общие замечания:

1. В разделе «материалы и методы» приведены методы, для которых не понятно из текста диссертации, где и как они применялись, например «Сортировка клеток, активированная флуоресценцией (FACS)», «Объединения перекрывающихся

фрагментов ДНК с помощью ПЦР», «Анализ люминесценции», «Обработка пДНК экзонуклеазой T5».

2. Зачем в работе использовались компетентные клетки, приготовленные по двум протоколам - для электропорации и для хитшока? Какая у них была эффективность трансформации?
3. Для каждого «повтора образца» в среднем было получено ~1.5 млн одноконцевых прочтений. Получается, что, например для библиотеки «41-48» (с числом мутаций 53837) на одну мутацию в среднем приходится около 28 прочтений. Кажется, что такая глубина секвенирования библиотек может не позволить адекватно количественно оценить, особенно небольшие, изменения в экспрессии исследуемых мутаций и привести к статистически ненадежным выводам.
4. При анализе экспрессии методом ОТ-ПЦР после трансфекции библиотекой плазмидной ДНК, существует риск ко-амплификации остаточной трансфицирующей плазмидной ДНК, которая может присутствовать в образце тотальной РНК даже после обработки ДНКазой. Следует пояснить, проводился ли такой контроль экспериментов и каковы его результаты.
5. В качестве контроля для нормирования использовалась исходная плазмидная библиотека. Более релевантным представляется использование в качестве нормирующего образца плазмидной ДНК, выделенной из клеток после трансфекции, так как это позволяет учесть систематические погрешности, вносимые этапами трансфекции и внутриклеточной репликации плазмид. Необходимо пояснить, почему был выбран менее строгий подход.
6. Представляется методологически более строгим первоначально проверить общую гипотезу о наличии функциональных регионов в изучаемой области с помощью грубого мутагенеза (например, делеционного анализа). После идентификации таких регионов проведение в них прицельного насыщающего мутагенеза MPRA позволило бы получить максимально детализированную функциональную карту с высокой эффективностью использования ресурсов, избежав затрат на скрининг заведомо нерелевантных участков.
7. Почему после трансфекции библиотек не проводили селекцию клеток по флуоресценции eGFP? Это позволило бы существенно обогатить пул последовательностей, оказывающих влияние на экспрессию гена.
8. Использованный подход (вставка рандомизированных последовательностей) по своей природе оценивает не функцию нативной последовательности, а регуляторный потенциал новой последовательности в данном локусе. Поэтому вывод о влиянии именно расположенных там последовательностей на экспрессию eGFP является методологически некорректным. Для таких утверждений необходимо провести анализ однонуклеотидных замен в существующем контексте.
9. Интересной особенностью эксперимента является ожидаемое наличие в каждой библиотеке одинаковых мутантных последовательностей, которые по сути вставлены в разные позиции. Это создаёт прямую основу для систематического анализа влияния местоположения (позиционного контекста) на регуляторную функцию одной и той же последовательности. Было бы интересно получить и обсудить эти результаты на полученных данных, так как сравнение активности идентичных последовательностей в разных точках (+17, +29, +41, +52 п.н.) может дать ценные данные о том, как

локальное окружение (например, соседние с СПА элементы, структура РНК) модулирует их эффект.

Заключение

Диссертация Летягиной Анны Евгеньевны «Анализ влияния последовательности, расположенной после сайта полиаденилирования, на уровень зрелой мРНК репортёрного гена eGFP в культивируемых клетках человека НЕК293Т», представленная на соискание ученой степени кандидата биологических наук по специальности 1.5.7 – «Генетика», оформлена в соответствии с требованиями научного цитирования, библиографический список составлен корректно, содержание работы соответствует паспорту специальности 1.5.7 – «Генетика». Основное содержание диссертации отражено в 3 научных публикациях автора в изданиях, рекомендованных для публикации результатов диссертационных исследований по специальности 1.5.7 – «Генетика». Диссертационная работа Летягиной А. Е. является самостоятельной, законченной научно-квалификационной работой. По критериям актуальности, научной новизны и практической значимости результатов диссертация соответствует требованиям пункта п.9 «Положения о присуждении ученых степеней», утвержденного Постановлением Правительства РФ «О порядке присуждения ученых степеней» от 21 сентября 2013 г. N 842 (в ред. от 01.01.2025 г.), предъявляемым к кандидатским диссертациям. Автор работы, Летягина Анна Евгеньевна, заслуживает присуждения ученой степени кандидата биологических наук по специальности 1.5.7 – «Генетика».

Заведующий отделом функциональной геномики Федерального государственного бюджетного научного учреждения «Медико-генетический научный центр имени академика Н.П. Бочкова», доктор биологических наук (1.5.7 – генетика).

Скоблов Михаил Юрьевич

Контактная информация:

Адрес: 115522, Москва, ул. Москворечье, д. 1

Тел.: +7 (499) 612-80-45,

e-mail: mskoblov@gmail.com

16 января 2026 года

Подпись доктора биологических наук Скоблова Михаила Юрьевича заверяю.

Ученый секретарь Федерального государственного бюджетного научного учреждения «Медико-генетический научный центр имени академика Н.П. Бочкова», кандидат медицинских наук



Воронина Екатерина Сергеевна