

Отзыв официального оппонента на диссертацию  
**Червонцевой Зои Сергеевны**  
**«Влияние вторичной структуры мРНК на экспрессию генов»,**  
представленную на соискание ученой степени кандидата биологических наук  
по специальности 1.5.8 – математическая биология, биоинформатика

**Актуальность темы**

Вторичная структура мРНК оказывает большое влияние на экспрессию генов, во многом обуславливая эффективность трансляции, время жизни и внутриклеточную локализацию транскриптов, а также внося существенный вклад в специфичность узнавания мРНК модифицирующими ферментами. При этом связь вторичной структуры в начале кодирующей части с эффективностью трансляции мРНК у бактерий, а также зависимость частоты редактирования аденинов в мРНК головоногих моллюсков от структуры транскриптов изучены недостаточно.

**Основные научные результаты и общая характеристика работы**

В диссертации описаны результаты выполнения нескольких проектов, объединённых общей методической базой и пересекающихся по научной тематике. В первой части работы проанализированы экспериментальные данные по влиянию нескольких десятков тысяч вариантов 30-нуклеотидных вставок в начале кодирующей части гена флуоресцентного белка на эффективность его экспрессии в клетках *E. coli*. Во второй проведено сравнение склонности к образованию вторичных структур у мРНК, кодирующих эквимоллярные субъединицы мультимерных белковых комплексов, закодированные в одном опероне. Далее изучена связь стабильности вторичной структуры мРНК *E. coli* со скоростью их деградации и сделана попытка проследить связь структуры с паттернами эволюции. Наконец, показана связь стабильности вторичной структуры со степенью редактирования аденинов в матричных РНК мягкотелых головоногих моллюсков и оценена её роль в скоррелированном редактировании пар аденинов.

Результаты изложены на 100 страницах, диссертация включает 28 рисунков и 3 таблицы, список литературы содержит 119 наименований. Структура диссертации несколько отличается от стандартной: за обзором литературы следуют три главы, соответствующие вышеупомянутым научным темам работы, в каждой из которых есть собственные разделы «Материалы и методы» и «Результаты и обсуждение». На мой взгляд, с учётом сложной структуры работы такое построение является оправданным, т.к. помогает читателю сопоставить методики, применённые в каждом из разделов, с полученными результатами. Кроме основных разделов (включающих также «Заключение» и «Выводы») и вспомогательных частей, работа содержит приложения с информацией о нуклеотидных последовательностях и штаммах *E. coli*.

**Научная новизна и значимость работы**

Все представленные в работе данные получены впервые, опубликованы в рецензируемых научных журналах и доложены на отечественных и международных конференциях. Результаты могут иметь значение в биотехнологии для оптимизации продукции рекомбинантных белков и в медицинской генетике для оценки влияния мутаций в некодирующих областях генов, а также быть использованы при чтении курсов по молекулярной биологии, системной биологии и эволюционной генетике в высших учебных заведениях.

## **Достоверность полученных результатов, обоснованность положений и выводов**

Выносимые на защиту научные положения основаны на анализе обширного материала, в том числе (в первой части исследования) оригинальных данных, полученных партнёрами по проекту. Методики применены корректно, использованы сложные статистические тесты и критерии. Достоверность полученных результатов и сделанных выводов не вызывает сомнений. Выводы соответствуют поставленным задачам. Качество работы подтверждается также и тем, что основные результаты опубликованы в виде трёх статей в рецензируемых научных журналах, в одной из которых (в высокорейтинговом журнале *Nucleic Acids Res*) соискатель выступила со-первым автором.

## **Замечания к работе**

Несмотря на хорошее качество работы, нельзя не отметить ряд недостатков, касающихся как оформления, так и сути работы. Так, несмотря на хороший в целом язык, на грамотное, стилистически выдержанное изложение и на несущественное количество опечаток, в тексте встречается изрядное количество жаргонизмов, некорректных или просто неудачных выражений. В качестве примеров можно привести: «работы репортируют», «малой единицы рибосомы», «элонгации трансляции», «пептидил-сайт», «пептиды... на 5'-концах транспортируемых белков», «шпильки, закодированные в концах транскриптов», «матричных РНК генов,... закодированных в одном опероне» (в опероне закодирована одна мРНК, несколько белков и не «закодировано» ни одного гена), «дезаминирование осуществляют деаминазы», «спаренных позициях генов» и другие.

Обзор литературы, пожалуй, излишне лаконичен для кандидатской диссертации (всего 12 страниц полуторным интервалом) и не содержит ни одной иллюстрации! Теме редактирования РНК и вовсе посвящено всего 1.5 страницы, чего может оказаться недостаточно для того, чтобы снабдить неподготовленного читателя всей информацией о структурных детерминантах редактирования, необходимой для последующей оценки новизны полученных результатов. При этом из 100 страниц диссертации 4 занимает оглавление, ещё 4 – список рисунков и таблиц, а собственно содержательная часть (обзор литературы и изложение результатов) занимает чуть больше половины объёма.

В работе также много других досадных мелочей, как-то: разный формат ссылок в списке литературы (где-то перечень авторов стоит в начале ссылки, где-то в конце, а где-то и вовсе отсутствует); некоторая небрежность в подписях к рисункам (в ряде случаев отсутствует общее название рисунка: описание начинается с легенды к панели «а»; порой без обращения к основному тексту нельзя понять, что там изображено: например, из Рис. 2.14б неясно, о каких конструкциях там идёт речь: кривые зашифрованы и читателю приходится догадываться, что нужно найти последовательности с таким же шифром в Приложении); путаница с нумерацией рисунков в разделе 2.2.5 (и вообще непонятно, почему Рис. 2.8г и 2.8д, явно относящиеся к этой главе, расположены на Рис. 2.8, посвящённом другим вещам, а не на Рис. 2.11); неидеальное соответствие заголовков содержимому разделов (например, в методическом разделе с названием «Получение плазмид, содержащих ген белка CER с рандомизированными вставками» описан дизайн всей «мокрой» части эксперимента, включая «трансфекцию» (!) клеток *E. coli* (очевидно, автор имела в виду трансформацию), сортировку клеток и прочее); иногда – упоминание нетривиальных фактов без должного разъяснения (например, пассаж про касугамицин на стр. 20 – кстати, в этом предложении сразу две опечатки).

В некоторых местах автору, на мой взгляд, следовало быть более аккуратной в формулировках и выводах. Например, заявление о том, что в работе «Впервые исследовано влияние случайных вставок в начале гена на эффективность инициации трансляции мРНК *Escherichia coli*» едва ли можно признать корректным с учётом наличия нескольких весьма известных работ, где был использован похожий подход (Goodman et al, 2013 в *Science*; Kudla et al., 2009 в том же журнале; другие статьи, в том числе цитируемые автором), а также с учётом того, что все выводы в этой части работы делали на основе оценки уровня репортерного белка, а он мог зависеть отнюдь не только от инициации трансляции, но и от стабильности белка, мРНК и ряда других причин, и к тому же сами авторы связывают наблюдаемые эффекты не столько с инициацией трансляции, сколько с элонгацией. Вообще, вышеупомянутые статьи (а также другие, опубликованные в 1990-х годах группой van Duin), казалось бы, были во многом посвящены вопросам, сильно перекрывающимся с теми, которые ставили перед собой авторы в первой части исследования: влиянию редких кодонов и вторичной структуры в начале кодирующей части гена на эффективность трансляции. Причём выводы там были сделаны отчасти похожие. Настоящая работа наверняка добавила очень существенные нюансы, однако объяснения этих нюансов и развёрнутого сравнения полученных результатов с данными предыдущих работ (которое обычно принято излагать в разделе «Обсуждение результатов») мне, к сожалению, найти не удалось.

Утверждение «Один и тот же фрагмент чаще всего попадал в ту же фракцию, а если и оказывался в другой, то чаще всего в соседней (Рисунок 2.2)» (стр. 31) вызывает сомнения, поскольку из Рис. 2.2 следует, что это неверно для двух из пяти фракций (крайних, 1 и 5) – и это, наверное, следовало бы отметить.

Вывод о том, что вставки из эффективно транслируемых фракций обеднены ШД-подобными участками (стр. 34), кажется, не стоит напрямую соотносить с результатами работ [59] и [66] (фраза «Это подтверждает предыдущие наблюдения о негативном влиянии внутригенных ШД-подобных участков на эффективность трансляции»), поскольку в тех работах речь шла о «притормаживании» элонгирующей рибосомы, а тут – скорее всего, о помехах инициации трансляции: ясно, что непродуктивное связывание 30S субчастицы рибосомы в окрестности инициаторной области (при отсутствии вблизи AUG-кодона в рамке) будет сильно снижать эффективность инициации с помощью аутентичной пары ШД-AUG, но вовсе не из-за эффектов на элонгацию, а просто из-за создания стерических затруднений для посадки второй рибосомы на «правильный» AUG.

Местами из-за недостаточных разъяснений у читателя возникают вопросы, которые хочется задать автору. Например, почему «Сила эффекта аминокислоты была посчитана как взвешенная сумма эффектов кодонов, соответствующих этой аминокислоте, с весами, соответствующими частотам этих кодонов в генах *E. coli*» (стр. 29)? Учитывались ли в анализе влияния вставок на уровень продукта трансляции наличие во вставке AUG-подобных кодонов, особенно не в рамке со стартовым AUG? При прочтении раздела 3.2.1, где представлены данные о корреляции степени структурированности участков полицистронных мРНК, кодирующих эквимоллярные субъединицы одного комплекса, хочется задать вопрос о количестве проанализированных пар: если судить по числу точек на Рис. 3.1а и 3.1в, пар было совсем немного, всего 15 и 8 штук в двух выборках, так ли это?

Хочется также задать вопрос по поводу диаграммы на Рис. 2.10: здесь неясно, как интерпретировать высоту столбиков, представляющих эффективность трансляции (оси с легендой нет, подпись неинформативна, а контроль без второго AUG в опыте отсутствует).

Таким образом, ответить на вопрос, поставленный в начале описания этого опыта («Действительно ли дополнительный старт-кодон AUG оказывает положительное влияние на трансляцию?») на основании этого рисунка затруднительно. На мой взгляд, результат с AUG, расположенными неподалёку от старта, довольно тривиален, а вот почему многие удалённые AUG также оказывали влияние (при том, что на нужном расстоянии от них не было последовательности Шайн-Далгарно), было бы, наверное, полезно обсудить. Кстати, дизайн участков, в которые вставляли AUG, мне также остался непонятен: судя по Приложению, во всех конструкциях последовательности похожи, но начиная с конструкции «+11», в них начинает варьировать количество С на 5'-конце, из-за чего кодоны и кодируемые ими аминокислоты оказываются разными. В итоге автор сравнивает между собой конструкции, различающиеся по нескольким параметрам. Делать какие-то однозначные выводы при таком подходе, мне кажется, довольно рискованно. Возможно, здесь есть какой-то скрытый смысл, но в тексте диссертации он, к сожалению, никак не отражён.

В главе 3.2.2, повествующей об отсутствии явной связи между общей структурированностью мРНК и скоростью ее деградации, при анализе по 100-нуклеотидным фрагментам выделены группы генов с положительной (211 генов) и с отрицательной (161 ген) корреляциями между уровнем структурированности и временем жизни фрагмента – однако больше об этих группах ничего не сказано. Была ли попытка выявить какие-то общие черты внутри групп и – главное – различающиеся черты между группами? Всё-таки сам анализ корреляции едва ли является достойной самоцелью: всегда хочется понять, стоят ли за находками какие-то механизмы (если же анализ не проводили, считая заранее известным, что за корреляциями ничего не стоит, то зачем тогда их искали?)

В главе 3.2.2 данные об эффективности трансляции, вторичной структуре и скорости деградации брали из разных источников – в таких случаях, возможно, имело смысл уточнить, насколько одинаковыми были штаммы, среда и условия роста клеток, в которых были получены данные разными группами авторов.

Также в названии раздела 4.2.6 и в Выводе 8 («Сайты, сближенные благодаря структуре, чаще редактируются одновременно») слово «одновременно», как мне кажется, немного вводит в заблуждение: на самом деле показать использованными в работе методами, что сайты редактируются именно «одновременно», нельзя – но можно говорить об их скоординированном редактировании.

Наконец, вызывает некоторое сожаление, что иногда автор не утруждает себя объяснить, зачем нужно было делать то или иное сравнение или искать корреляцию – какую цель преследовала автор, на какой вопрос хотела ответить. Из-за этого иногда создаётся впечатление, что часть работы была задумана по принципу «А давайте посмотрим, нет ли тут корреляции – и если есть, то придумаем, как её объяснить». Например, глава 3.2.3 начинается с предложения «Далее мы посмотрели, как связана структура мРНК и эволюция последовательности» - но зачем это делали, не объяснено. Ситуация усугубляется тем, что сделанная находка, по мнению самого автора, скорее всего, является артефактом – на чём эта глава объёмом в полстраницы благополучно заканчивается. Кстати, замены С на U в контексте структуры мРНК имеют, пожалуй, особый интерес, т.к. такая замена с меньшей вероятностью изменяет вторичную структуру: и С, и U в составе транскрипта способны спариваться с G.

Несмотря на все эти незначительные недочёты и оставшиеся неосвещёнными небольшие второстепенные вопросы, в работе получены и адекватным образом изложены бесспорно очень интересные результаты, которые представляют большую ценность для науки. Особенно

интересна и оригинальна, на мой взгляд, обнаруженная зависимость от метаболической стоимости аминокислот, а также выявленные в последней части работы закономерности в структурной организации областей мРНК, окружающих редактируемые сайты.

### **Заключение**

На основании всего вышеизложенного можно заключить, что диссертационная работа Червонцевой Зои Сергеевны «Влияние вторичной структуры мРНК на экспрессию генов», выполненная в Институте проблем передачи информации им. А.А. Харкевича Российской академии наук, представляет собой цельное и законченное исследование. Работа выполнена на высоком научном и методическом уровне, соответствует критериям, предъявляемым к кандидатским диссертациям, и соответствует специальности 1.5.8. – математическая биология, биоинформатика.

Автореферат достаточно полно отражает содержание диссертации. Основные результаты опубликованы в трех статьях в международных журналах, входящих в международные библиографические базы данных, и доложены на профильных научных конференциях. В диссертации отражен личный вклад автора в указанные публикации.

По актуальности, научной новизне, научно-практической значимости диссертация З.С. Червонцевой «Влияние вторичной структуры мРНК на экспрессию генов» является законченной научно-квалификационной работой и отвечает требованиям п. 9 «Положения о порядке присуждения ученых степеней», утвержденным постановлением Правительства Российской Федерации № 842 от 24 сентября 2013 г., а ее автор З.С. Червонцева заслуживает присуждения искомой степени кандидата биологических наук по специальности 1.5.8 – математическая биология, биоинформатика.

### **Официальный оппонент:**

Дмитриев Сергей Евгеньевич,  
кандидат биологических наук,  
заведующий отделом взаимодействия вирусов с клеткой  
НИИ физико-химической биологии имени А.Н.Белозерского  
Московского государственного университета имени М.В.Ломоносова,  
Адрес: 119234 Москва, Ленинские горы, 1с40  
Эл. почта: [sergey.dmitriev@belozersky.msu.ru](mailto:sergey.dmitriev@belozersky.msu.ru)

Специальность, по которой официальным оппонентом была защищена диссертация:  
03.00.03 – молекулярная биология

**Подпись Дмитриева С.Е. заверяю**

Учёный секретарь НИИ ФХБ имени А.Н.Белозерского МГУ,  
к.б.н.



И.А. Севостьянова

05.09.2023