

ОТЗЫВ

официального оппонента доктора биологических наук член-корреспондента РАН
А. В. Кочетова о диссертационной работе Денисова Степана Владимировича на
тему «ОТБОР И ЭПИСТАЗ В САЙТАХ СПЛАЙСИНГА», представленной на
соискание ученой степени кандидата биологических наук по специальности
03.01.09 – математическая биология, биоинформатика

Диссертация С.В. Денисова посвящена изучению эволюции сайтов сплайсинга в структуре эукариотических генов. Сплайсинг является одним из фундаментальных механизмов, определяющих особенности хранения и реализации генетической информации. Информация о локализации конститутивных и альтернативных сайтов сплайсинга также необходима для предсказания полного кодирующего потенциала эукариотических генов. Работа С.В. Денисова посвящена исследованию эволюции этого класса регуляторных последовательностей, что важно как для развития существующих представлений об их возникновении, так и для выявления особенностей их структурно-функциональной организации. Несмотря на то, что исследование сайтов сплайсинга и их эволюционной истории традиционно привлекает большое внимание, ряд ключевых проблем по-прежнему нельзя назвать решенными. Это позволяет определить тему диссертационной работы как актуальную и представляющую большой интерес для эволюционной геномики и биологии в целом.

Диссертация С.В. Денисова имеет общий объем в 166 страниц, включая 27 рисунков и 5 таблиц, построена по следующей схеме: «Введение» (7 стр.), «Обзор литературы» (59 стр.), «Данные и методы» (11 стр.), «Результаты и обсуждение» (54 стр.), «Основные результаты и выводы» (2 стр.), «Приложения» (10 стр., включают 11 рисунков и 2 таблицы), отдельно в конце работы вынесен раздел «Благодарности» (1 стр.). Список литературы содержит 218 источников, все в международных изданиях на английском языке. Результаты работы были представлены на авторитетных отечественных и международных научных конференциях. По теме диссертации С.В. Денисов опубликовал 3 статьи в авторитетных международных научных журналах, в которых он является первым автором.

В разделе «Введение» автор обосновал актуальность своей работы, обозначив ее цель как «изучить, каким образом естественный отбор, генетический дрейф и эпистаз между позициями влияют на эволюцию сайтов сплайсинга, а также изучить консервативность регуляции кассетных экзонов на примере цис-регуляторного

элемента UGCAUG». Для достижения поставленной цели сформулированы 7 задач, связанных с использованием конкретных выборок и моделей. С моей точки зрения цель и задачи сформулированы в целом адекватно, хотя некоторые термины описаны в тексте диссертации позднее, что затрудняет понимание при последовательном прочтении работы (например, «сила сайта сплайсинга», «сила нуклеотида» могут быть выражены различными способами).

Некоторые аспекты оформления диссертационной работы С.В. Денисова отклоняются от общепринятых. Например, в разделе автореферата «Общая характеристика работы» нет пункта «Положения, выносимые на защиту». Однако, в тексте диссертационной работы в разделе «Введение» есть пункт «Основные результаты и положения, выносимые на защиту». В автореферате есть раздел «Степень достоверности и апробация результатов», в котором приведены данные о публикациях по материалам работы. По видимому, более правильно использовать «Апробация результатов диссертационной работы».

Замечания к разделу «Введение»: Есть несколько неудачных выражений:

Стр. 4 «...структура и экспрессия мРНК генов человека...» - есть устойчивые выражения «уровень экспрессии гена», «содержание мРНК», «уровень транскрипции гена», выражение «экспрессия мРНК» лучше не использовать.

Стр. 9 Опечатка в слове «Проанализировали».

Глава «Обзор литературы» содержит три раздела. В первом («Альтернативный сплайсинг») приведено краткое описание собственно базовых процессов сплайсинга, структуры сайтов и некоторые данные о регуляции альтернативного сплайсинга. Во втором разделе приведен интересный обзор известных данных и моделей эволюции сплайсинга. В третьем разделе «Отбор и эпистаз» приведен краткий обзор современных представлений о молекулярной эволюции в приложении к изучаемому объекту. Хочу отметить, что обзор литературы написан в ярко выраженном авторском ключе и читать было очень интересно. Хорошее впечатление произвело описание методических подходов, применяемых для моделирования и статистических оценок вероятности тех или иных классов молекулярных эволюционных событий. В целом, существенных замечаний к этому разделу нет. Тематика работы фундаментальная и полномасштабный обзор в одной работе сделать затруднительно, но автору удалось скомпилировать информацию таким образом, что получилось вполне достаточное описание существующего положения дел в соответствующей области науки.

Замечания по главе «Обзор литературы». Есть некоторое (небольшое) количество ошибок и грамматических небрежностей (стр. 25 «Они высоко консервативны белков среди многоклеточных», стр. 44 «К группе похожих на S.

cerevisiae относятся *S. cerevisiae*, ...)», неудачных выражений (стр. 48 «эффективная численность популяции варьирует вдоль генома», стр 53 «длинный гаплотип») и т.п.

Глава «Данные и методы» содержит два раздела. В первом приведена информация о выборках нуклеотидных последовательностей, во втором описаны использованные в работе методы. Описание достаточно подробное.

Работа выполнена в области эволюционной геномики и в ней были поставлены задачи, для решения которых необходимо определение эволюционных расстояний и реконструкция предковых последовательностей. Объект исследования (сайты сплайсинга) имеет свою специфику – это короткие сайты, часть позиций в которых вырождена. Каким образом те или иные комбинации нуклеотидов в структуре донорных или акцепторных сайтов влияют на эффективность собственно процесса сплайсинга – в большинстве случаев сказать сложно и эта область активно изучается. Существуют определенные структурно-функциональные зависимости – например, консенсусная последовательность считается более функционально активной в силу комплементарности соответствующим малым ядерным РНК. Однако эффективной (предсказательной) модели РНК-белковых взаимодействий в процессе сплайсинга пока нет и точное предсказание локализации сайтов сплайсинга в генах млекопитающих *ab initio* не реализовано.

С учетом сложности задачи С.В. Денисов применил ряд подходов, основанных на достаточно простых приближениях. Степень сходства с консенсусом была обозначена как «сила сайта» или «сила нуклеотида» (по сути, это вариант информационного содержания). Следует отметить, что сила сайта может коррелировать с эффективностью сплайсинга, однако активность разных комбинаций нуклеотидов может в существенной степени различаться. Условность (приближенность) критерия достаточно очевидна (об этом говорит сама процедура определения консенсуса (стр. 81 – «один или два самых частых нуклеотида мы называли консенсусными..»)). Тем не менее, для оценки на больших выборках, сила сайта/сила нуклеотида вполне может использоваться. Мне представляется очень удачным применение метода реконструкции предковой последовательности, а также анализ ковариационных матриц для выявления потенциальных меж-позиционных функциональных связей.

Глава «Результаты и обсуждение» содержит три раздела: первый посвящен изучению положительного и отрицательного отбора в сайтах сплайсинга, второй – коррелированной эволюции позиций в этих сайтах, эти разделы связаны между собой методически и объектно. В третьем разделе (отбор в окрестностях сайтов

сплайсинга) представлены данные, полученные автором значительно раньше и на других выборках.

В ходе выполнения работы Степан Владимирович получил несколько новых интересных результатов, значимых для развития существующих представлений в области эволюционной геномики. Во-первых, было продемонстрировано действие слабого отбора на консенсусные и неконсенсусные нуклеотиды в позициях сайтов сплайсинга (закрепляющего и движущего, соответственно). Характерно, что на молодые сайты сплайсинга в генах человека действие отбора значительно более выражено, что дополнительно подтверждает выявленную закономерность. Проблема эффективности слабого отбора у видов с небольшим эффективным размером популяций обсуждается очень активно. Автором получено подтверждение одной из концепций, согласно которой такой отбор имеет место наряду с генетическим дрейфом. Эта тема является одной из фундаментальных проблем эволюционной биологии и исследования весьма востребованы, что позволяет оценить представленный в работе результат как высокозначимый. Во-вторых, в работе был проведен анализ эволюционной истории позиций в донорных и акцепторных сайтах сплайсинга, что позволило сделать ряд наблюдений о тонкой структуре сайтов, межпозиционных взаимодействиях, роли отбора против определенных динуклеотидов, в том числе в формировании полипиримидинового тракта. В частности, интересен результат об отрицательных корреляциях между силой нуклеотидов в экзонной и интронной частях сайтов. Автором выдвинут ряд оригинальных предположений о механизмах, лежащих в основе этих закономерностей. Следует учесть, что автор рассматривал выборки донорных и акцепторных сайтов сплайсинга независимо: было бы интересно посмотреть наличие взаимосвязей между парными донорными и акцепторными сайтами. Эта часть работы также представляет собой хороший задел для планирования экспериментов.

Замечания к главе «Результаты и обсуждение».

На рис. 14 д-з синие столбики ошибочно названы красными. Автору следует определиться, какое склонение русскоязычного названия *Macaca mulatta* использовать: макак (в автореферате) или макака (в диссертационной работе, например – стр. 96). При выборе в качестве нейтрального контроля нуклеотидной последовательности интрона, следует иметь в виду, что композиционный состав некодирующих последовательностей на самом деле довольно разный – и по частотам мононуклеотидов, и по предпочтительному использованию динуклеотидов. Такие районы, как 5'-НТП, 3'-НТП и интроны могут отличаться

достаточно сильно и быть нейтральными контролями друг для друга не всегда могут.

Глава «*Основные результаты и выводы*» на самом деле представляет собой просто выводы, правда структурированные на три группы. В целом, они адекватны полученным результатам, хотя в некоторых случаях формулировки представляются излишне категоричными (например, «отбор против динуклеотида AG есть основная причина ковариаций...»).

Заключение: В целом, работа производит впечатление хорошего фундаментального исследования, внесшего вклад в решение важной научной проблемы. Это законченная научно-квалификационная работа, новизна и значимость которой не вызывают сомнений. Полученные результаты являются новыми и выводы диссертации вытекают из полученных результатов. Автореферат отражает содержание работы. Материал диссертации соответствует указанной специальности. Диссертация апробирована на международных конференциях, результаты опубликованы в авторитетных международных научных журналах. Таким образом, диссертационная работа С.В. Денисова полностью соответствует требованиям ВАК, предъявляемым к кандидатским диссертациям (пп. 9-14 «Положения о присуждении ученых степеней» за №842 от 245 сентября 2013 г.), а ее автор заслуживает присуждения искомой степени кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика.

Кочетов Алексей Владимирович

доктор биологических наук (03.02.07- генетика,
03.01.09 - математическая биология, биоинформатика)
чл.-корр. РАН

заведующий лабораторией генной инженерии,
заместитель директора по научной работе
ФГБНУ «Федеральный исследовательский
центр Институт цитологии и генетики
Сибирского отделения Российской академии наук»
(ИЦиГ СО РАН)

630090 г. Новосибирск пр. академика Лаврентьева, 10

Тел. 8383-3634994

ak@bionet.nsc.ru

03.04.2017

