

Отзыв официального оппонента

на диссертационную работу Денисова Степана Владимировича “Отбор и эпистаз в сайтах сплайсинга”, представленную на соискание ученой степени кандидата биологических наук по специальности 03.01.09 – “математическая биология, биоинформатика”

Диссертационная работа С.В. Денисова посвящена изучению эволюции сайтов сплайсинга млекопитающих и мух рода *Drosophila*. Сплайсинг – процесс вырезания интронов и сшивки экзонов – играет ключевую роль в формировании белкового разнообразия у высших эукариот. На границах экзонов и интронов располагаются специальные последовательности – сайты сплайсинга – они распознаются сплайсосомой, обеспечивающей точное распознавание границ экзонов и интронов и последующее вырезание интронов. В каждой позиции сайта сплайсинга нуклеотиды делятся на консенсусные, т.е. часто встречающиеся в данной позиции у разных сайтов сплайсинга, и неконсенсусные (редкие). В диссертационной работе С.В. Денисова исследовано, как естественный отбор, генетический дрейф и эпистаз влияют на эволюционные изменения, происходящие в сайтах сплайсинга. За последние 10-15 лет накоплен огромный объем данных качественно аннотированных сайтов сплайсинга в геномах многих видов млекопитающих, а так же модельных организмов, таких как *Drosophila*, что позволяет изучать эволюцию индивидуальных позиций сайтов сплайсинга.

В диссертационной работе С.В. Денисова были исследованы следующие аспекты эволюции сайтов сплайсинга: (1) влияние естественного отбора на консенсусные и неконсенсусные нуклеотиды и генетический дрейф, отдельно исследованы молодые сайты сплайсинга (то есть появившиеся на линии человека); (2) влияние замен в одних позициях сайтов сплайсинга на замены в других позициях, рассмотрены вероятные причины, обуславливающие это влияние; (3) отдельно исследована консервативность системы регуляции альтернативного сплайсинга на основе цис-регулятора UGCAUG.

Диссертация изложена на 166 страницах. Текст работы разбит на восемь разделов: введение, обзор литературы, данные и методы, результаты и обсуждение, основные результаты и выводы, благодарности, приложение и список литературы. Работа содержит 27 рисунков и 5 таблиц. Список литературы содержит 218 наименований. Приложение содержит 11 рисунков и 2 таблицы.

Во введении кратко сформулированы основные положения, выносимые на защиту, а также прояснена актуальность работы.

Обзор литературы изложен на 59 страницах, состоит из трёх частей. В первой части рассматривается молекулярный механизм сплайсинга и вопросы регуляции альтернативного сплайсинга. Вторая часть посвящена известным аспектам микро- и макроэволюции сплайсинга. В третьей части рассматриваются важные для дальнейшего изложения фундаментальные эволюционно-генетические понятия, такие как положительный и отрицательный отбор, эпистаз и эпистатический отбор. Также рассматриваются статистические методы определения положительного и отрицательного

отбора на основании геномных данных. Литературный обзор изложен понятным языком, обильно иллюстрирован и снабжен достаточным количеством ссылок на оригинальные научные статьи.

В разделе “Данные и методы” указывается, что автор диссертации в своей работе использовал разнообразные публично доступные данные: геномные последовательности, аннотацию генов, транскриптомы, данные по внутривидовой изменчивости (полиморфизмы) и межвидовой дивергенции (множественные выравнивания), экспрессионные данные, а также данные по рекомбинации. Используемые методы излагаются достаточно подробно. Часть методов, разработанных лично автором, диссертации изложены в разделе “Результаты и обсуждение”.

В разделе “Результаты и обсуждение” излагаются результаты работы соискателя. Он состоит из трех частей. В первой части рассматривается действие дрейфа и отбора на эволюцию сайтов сплайсинга. Для детектирования положительного и отрицательного отбора используется разработанный автором тест, являющийся аналогом теста D_n/D_s для некодирующих функциональных элементов (а именно, для сайтов сплайсинга). Применение этого теста представляется уместным и корректно реализованным. Было показано, что на неконсенсусные нуклеотиды действует положительный отбор, в то время как консенсусные нуклеотиды защищены отрицательным отбором. Автор оценивает силу отбора, действующего на консенсусные и неконсенсусные нуклеотиды и приходит к выводу, что коэффициент отбора s сопоставим по порядку величины с $1/N_e$, где N_e – эффективный размер популяции. Другими словами, на сайты сплайсинга воздействует как генетический дрейф, так и естественный отбор и степень этого воздействия сопоставима. Таким образом, переходы из консенсусных нуклеотидов в неконсенсусные являются слабовредными, а обратные переходы – слабополезными. Данная работа, насколько мне известно, является первой, где прямо показано наличие слабополезных мутаций в сайтах сплайсинга. Интересно, что слабовредные и, главное, слабополезные замены происходят при неизменном ландшафте приспособленности, т.к. компоненты сплайсосомы неизменны на рассматриваемых эволюционных расстояниях. Это отличается от классической дарвиновской ситуации, где положительный отбор происходит при изменении ландшафта приспособленности. Этот результат представляется важным в фундаментальном аспекте. Кроме того было продемонстрировано, что на молодые в эволюционном отношении сайты действует гораздо более сильный положительный отбор, что как раз свидетельствует о резком изменении ландшафта приспособленности при возникновении нового сайта из интронных (некодирующих) последовательностей. Хочется надеяться, что автор в будущем изучит последнее наблюдение более подробно в контексте изменения экзон-интронной структуры генов. Показано также, что на неконсенсусные нуклеотиды некоторых сайтов сплайсинга действует сайт-специфический отрицательный отбор. Важно, что этот отбор не является отбором на сохранение сайтов слабыми, как предполагалось в некоторых более ранних работах. Хочется отметить оригинальность метода с помощью которого это было продемонстрировано.

Большинство сделанных в работе выводов подтверждены независимо на двух несвязанных эволюционных линиях: линии *Homo sapiens* после дивергенции от *Macaca mulatta* и *Drosophila melanogaster* после отхождения от *Drosophila simulans*, что позволяет предполагать их универсальность.

Во второй части раздела "Результаты и обсуждение" рассматривается другой любопытный феномен молекулярной эволюции: зависимость эволюции одних локусов от других. Автор рассматривал ковариации сил нуклеотидов, занимающих различные позиции внутри сайтов сплайсинга (сила нуклеотида в позиции отражает его совпадение с консенсусом). В результате критического анализа множества различных факторов автор делает вывод, что данные корреляции вызваны эпистазом, а существующий эпистатический отбор, приводит к усилению этих корреляций в ходе эволюции. Найденный периодический характер корреляций в акцепторных сайтах сплайсинга объясняется отбором против динуклеотида AG, который, как отмечается автором, является вырожденным случаем эпистатического отбора. Это наблюдение объяснимо с биологической точки зрения, т.к. по сути отражает отбор против конкурирующих сайтов сплайсинга, способных нарушить экзон-интронную структуру.

В третьем разделе рассматривает консервативность цис-регулятора сплайсинга UGCAUG. Исследуется выборка кассетных экзонов, экспрессирующихся преимущественно в мозге, для которой показано, что в интронах, следующих за этими экзонами последовательность UGCAUG встречается существенно чаще, чем ожидается. Анализ консервативности этих элементов позволил, предположить какие из них являются функциональными.

Несмотря на высокий уровень выполненной работы, текст диссертации не лишен недостатков. В тексте встречаются опечатки, которые, впрочем, не влияют на его восприятие. Более существенные замечания приведены ниже:

Стр. 99. "Сила положительного отбора на приобретение консенсусного нуклеотида сильнее у слабых акцепторных сайтов", кажется, что это противоречит данным на рис. 18Б?

Стр. 108, таб. 1. В позиции -3 в линии человека отклонение от ожидаемого значения изменения силы сайта равно 10.78 и не является, значимым, в тоже время, значительно меньшие отклонения в других позициях сайта в этой линии имеют высокую значимость. Как это объясняется?

Стр. 109, таб. 2. В позиции -4 в конститутивных экзонах наблюдается сильная потеря силы сайта в трех линиях, что означает отдаление сайта от консенсуса в этой позиции (эффект является значимым). Если верить рис. 2, то в этой позиции нет консенсуса, что тогда понимается по "силой" для этой позиции.

Раздел 4.1.5., стр. 95. Автор пишет: "метод парсимонии имеет тенденцию недооценивать частоту замен из неконсенсусных нуклеотидов в консенсусные, что и приводит к переоценке силы отрицательного и недооценки силы положительного отбора". Казалось бы, что метод парсимонии должен одинаково недооценивать, по сравнению с методом

макс. правдоподобия, частоты замен как из консенсусных, так и неконсенсусных нуклеотидов?

Стр. 80. На каком основании пренебрегаются ошибки в нейтральном контроле? Из общих соображений, величина ошибки в контроле не отличается от таковой в данных.

Раздел 4.1.4, стр. 89, представлены данные для средней межвидовой консервативности только для неконсенсусных нуклеотидов. Результаты по консенсусным нуклеотидам нигде не представлены.

Диссертационная работа С.В. Денисова является законченным самостоятельным исследованием, выполненном на высоком научном и методическом уровне. Значимость и достоверность результатов не вызывают сомнений. Сделанные выводы хорошо обоснованы. Результаты диссертации опубликованы в рецензируемых научных журналах и апробированы на международных и российских конференциях. Диссертационная работа С.В. Денисова полностью удовлетворяет требованиям пп. 9-14 "Положения о порядке присуждения ученых степеней", утверждённым Постановлением Правительства Российской Федерации №842 от 24 сентября 2013 г. (с изменениями в редакции Постановлений Правительства Российской Федерации №335 от 25 апреля 2016 г., №478 от 2 августа 2016 г.), предъявляемым к кандидатским диссертациям. С.В. Денисов безусловно заслуживает присуждения степени кандидата биологических наук по специальности 03.01.09 – "математическая биология, биоинформатика".

Подпись Неверова А.Д.
Ученый секретарь А.Д. Неверов заверяю
ФБУН ЦНИИ Эпидемиологии Роспотребнадзора

Неверов Алексей Дмитриевич,
кандидат биологических наук,
старший научный сотрудник,
Федеральное бюджетное учреждение
науки

Центральный НИИ Эпидемиологии
Роспотребнадзора

Email: neverov@pcr.ms

Телефон: (495) 672-11-58

06.04.2017

