

УТВЕРЖДАЮ

Зам. директора ИОГен РАН



О Т З Ы В

ведущей организации на диссертационную работу Храмеевой Екатерины Евгеньевны "Дальние взаимодействия в геномах эукариот и регуляция сплайсинга", представленную на соискание ученой степени кандидата биологических наук по специальности 03.01.09 – «математическая биология, биоинформатика».

Актуальность. Диссертация Храмеевой Екатерины Евгеньевны посвящена анализу организации и функционирования нуклеиновых кислот – РНК и ДНК. При этом основное внимание уделяется крупномасштабной структуре нуклеиновых кислот и механизмам, обеспечивающим пространственную укладку биомакромолекул. В случае РНК – это в первую очередь сплайсинг и механизмы, ответственные за удержание или вырезание конкретных инtronов. В случае ДНК – пространственная организация хромосом. Несмотря на кажущуюся пропасть между этими объектами, на самом деле диссертация является вполне органичным исследованием, в виду того, что такие процессы как транс-сплайсинг и формирование химерных транскриптов, чаще происходят на ко-экспрессирующихся и

пространственно сближенных участках ДНК, а главное, изучаются близкими методами, и их изучение осложняется практически идентичными методическими сложностями и артефактами.

Тема исследования Екатерины Евгеньевны находится в сердце фундаментальной молекулярной биологии. Актуальность этой темы связана в первую очередь с методами, появившимися в последние годы, обязанными своим происхождением внедрению технологий секвенирования нового поколения. В диссертации – это иммунопреципитация РНК и HiC. Благодаря использованию этих технологий появилась возможность получения информации о строении и взаимодействии макромолекул на уровне, представлявшемся невозможным еще каких-нибудь шесть лет назад. В настоящее время в фокусе исследований фундаментальные процессы молекулярной биологии. В то же время, пересмотр и углубление понимания фундаментальных молекулярно-биологических процессов в самой ближайшей перспективе приведет к практическим нововведениям в области биотехнологии и медицине, основанным на тонком понимании механизмов регуляции экспрессии и сплайсинга конкретных ключевых генов.

Работа Екатерины Евгеньевны здраво отражает захватывающий характер стадии, на которой находится наша наука. Среди представленных результатов некоторые имеют вполне фундаментальный характер и в том или ином виде войдут в учебники. В частности, к таким результатам относится корреляция пространственной организации хромосом и большого количества функциональных показателей, локально характеризующих хромосому — начиная от клеточной локализации продуктов экспрессии и кончая характеристиками хроматина и интенсивности экспрессии генов. В общем, работа Храмеевой Екатерины Евгеньевны лежит на переднем крае науки, имеет безусловную фундаментальную значимость, основана на

анализе данных, полученных самыми современными методами, и, таким образом, высокоактуальна.

Структура работы. Работа состоит из обзора литературы и трех глав. От классической компоновки диссертацию отличает отсутствие методической главы, а также то, что изложение и обсуждение результатов происходит в каждой главе отдельно. Такая структура обусловлена тем, что работа объединяет несколько исследований, отличающихся и объектом и методами. По сути — последняя глава в большой степени посвящена функциональной геномике, в то время как первые две главы относятся к транскриптомике. Обзор литературы дает прекрасное введение в биологию регуляции сплайсинга. Кроме того, в обзоре литературы излагаются детали экспериментальных методов, использованных для получения данных, и связанные с этими методами возможные артефакты. Глава «Поиск вторичных структур, участвующих в регуляции сплайсинга» содержит изложение метода поиска шпилек в РНК, предположительно участвующих в регуляции сплайсинга, с учетом комплементарности последовательностей в основании шпильки и консервативности соответствующих участков у ряда позвоночных. Во второй главе изложены результаты анализа данных CLIP-seq для белка hnRNPL и показана важная роль этого белка как в вырезании/удержании альтернативных экзонов, так и в сопряжении регуляции на уровне сплайсинга с регуляцией с помощью микро-РНК. Наконец третья глава, на мой взгляд самая интересная, содержит анализ химерных РНК в транскриптоме, с интерпретацией результатов как транс-сплайсинга или артефактов эксперимента. Диссертант проделала огромную работу по анализу необработанных данных, и в частности, сравнила результаты транс-сплайсинга с результатами хромосомной топологии, полученной с помощью методов HiC.

Каждая глава заканчивается коротким описанием полученных в этой главе результатов. Кроме того диссертация содержит заключение, в котором опять-таки кратко описываются полученные результаты и дается их оценка. Вообще, диссертация написана очень ясным языком и очень легко и с удовольствием читается.

Научная новизна. Автором впервые получен ряд замечательных результатов, которые хорошо изложены в выводах работы, поэтому я не буду их здесь полностью перечислять. Остановлюсь только на нескольких, на мой взгляд самых значительных достижениях. Во второй главе методом iCLIP, совмещенным с высокопроизводительным секвенированием, была построена точная полногеномная карта позиций связывания белка hnRNPL, в результате чего обнаружено большое количество позиций связывания в инtronах. Это результат ценный методически, поскольку отражает связывание с несплайсированным предшественником мРНК, объектом, который трудно изучать молекулярно-биологическими методами. Кроме того, сравнение участков связывания hnRNPL в окрестностях экзонов, включение которых активируется или репрессируется при нокауте hnRNPL, показало наличие многочисленных сайтов связывания hnRNPL в 3' районе интрана, только для репрессируемых сайтов сплайсинга. Таким образом, впервые удалось показать методом полногеномного анализа роль белка hnRNPL в участии альтернативного сплайсинга. Исследования дифференциальной экспрессии отдельных экзонов показали двойственную роль hnRNPL. Этот белок репрессирует включение экзона, если связывается в интронной области, в непосредственной близости от 3'-сайта сплайсинга, и, напротив, активирует включение, если связывается в интронной области вблизи 5'-сайта сплайсинга. Другой важный результат, полученный в этой же главе, — это существенно повышенная плотность участков связывания hnRNPL в окрестностях сайтов связывания микроРНК, что позволяет предположить

связь регуляции экспрессии на уровне миРНК и регуляции альтернативного сплайстига.

Красивейший результат, изложенный в третьей главе — высокое сходство различных функциональных характеристик хромосом, сближенных в пространстве согласно данным HiC. Обращает на себя внимание корреляция иерархии категорий генной онтологии «клеточный компонент», достигающая 0.98 по Спирмену.

Научная значимость результатов и их практическое значение.

Научная значимость результатов и практическое значение доказываются тем, что автором открыты новые явления в области строения участков мРНК, ответственных за альтернативный сплайсинг, а также ряд закономерностей, связанных с транс-сплайсингом и пространственной организацией хромосом. С точки зрения фундаментальной науки, открытые диссиденткой закономерности должны послужить отправной точкой для дальнейшего исследования механизмов регуляции экспрессии на разных уровнях, в частности связи сплайсинга с регуляцией на уровне микро-РНК. С научно-практической точки зрения открытые автором закономерности могут использоваться для понимания роли альтернативного сплайсинга при синтезе конкретных изоформ генов, значимых для развития различных заболеваний; известен ряд заболеваний, связанных с aberrantным сплайсингом. Кроме того в перспективе возможно повышение эффективности синтеза конкретных сплайсоформ путем модификации участка в ДНК окрестностях сайтов сплайсинга методами генной инженерии.

Замечания к работе. Работа в целом очень хорошо написана, однако имеет место ряд неясных моментов. Во-первых, большая часть главы 3 посвящена анализу артефактов при обработке транскриптомных данных. В частности, показано, что артефакты, по-видимому, связанные с конкретным оборудованием и пробоподготовкой в конкретных лабораториях, часто

перекрывают биологические эффекты. Мне, однако, не удалось выяснить из текста диссертации, производился ли контроль качества чтений (reads), взятых из баз данных, и фильтрация адапторов. Известно, что на заре секвенирования нового поколения оба этих процесса проводились недостаточно хорошо, и в настоящее время фильтрация ридов обязательна при использовании данных из SRA. Во-вторых, на рис. 2.2 (стр.63) показано существенное связывание белка hnRNPL в межгенниках (23.3%). Как интерпретировать эту цифру? Не может ли это быть связано с длинными некодирующими РНК? Это интересный момент, но он никак не обсуждается диссидентом. Там же, на стр. 66 обсуждается связывание hnRNPL в 5' инtronной области, активирующее включение экзона. Разница линий на картинке невелика, интересно, насколько эти данные статистически достоверны. Далее, на стр. 55 показано, что нарушение комплементарности участков E и F гена sf1 (рис 1.8А) приводит к вырезанию более длинного интрана из-за преимущественного использования внешнего акцепторного сайта сплайсинга вместо внутреннего акцепторного сайта. Создается впечатление, что разрушился внутренний акцепторный сайт. Не проверялось ли разрушением донорного сайта F работа всего механизма в целом?

Есть ряд и редакционных неточностей. На стр.25. В подписи к рис.11 упомянут комплекс Б, со сближенными сайтами сплайсинга, а на рисунке этот комплекс не указан. Названия u2AF и u2AF65 - это синонимы, и можно было бы выбрать один из них, а не оба. На стр. 30 сказано, что технология CLIP позволяет идентифицировать участки пре-мРНК, с которыми связывается белок, с точностью до одного нуклеотида. Между тем на стр. 57 для проверки повторяемости результатов CLIP брались интервалы от 5 до 30 нуклеотидов, то есть ожидаемая точность метода по-видимому ниже и это вполне учитывается автором. Очень трудно понять, где экзон, а где инtron на рисунке 2.3 — они никак не указаны и надо долго вчитываться в текст. На стр. 82 сказано, что ДНК-полимераза (фермент, который осуществляет

полимеразную цепную реакцию) также может ошибиться и перепутать участки генома, содержащие достаточно протяженные идентичные последовательности. На самом деле ДНК полимераза движется локально и “не знает” об идентичности участка генома. Видимо дело в каких-то артефактах PCR - например в гибридизации разных участков ДНК с идентичными последовательностями при отжиге.

Сделанные замечания не умаляют общего высокого уровня работы.

Рекомендации по практическому использованию. Результаты диссертации могут использоваться при проведении научных исследований в области молекулярной биологии и биоинформатики. Описанные в диссертации методы и статистические закономерности могут быть использованы в Федеральном государственном бюджетном учреждении науки Институт общей генетики им. Н.И. Вавилова РАН, Федеральном государственном бюджетном учреждении науки Институт молекулярной генетики РАН, Федеральном государственном бюджетном учреждении науки Институт математических проблем биологии РАН, Федеральном государственном бюджетном учреждении науки Институт биологии гена РАН, Федеральном государственном бюджетном учреждении науки Институт проблем передачи информации им. А.А. Харкевича РАН, Московском государственном университете имени М.В. Ломоносова, Новосибирском государственном университете и других учебных и научно-исследовательских организациях.

Заключение. Диссертационная работа Храмеевой Екатерины Евгеньевны «Дальние взаимодействия в геномах эукариот и регуляция сплайсинга» является законченным научно-квалификационным исследованием, в котором решена задача анализа определения механизмов регуляции альтернативного сплайсинга и транс-сплайсинга, осуществляемых как вторичной структурой РНК, так и регуляторными белками. Решение

этой задачи имеет большое значение для задач компьютерной геномики (пункт 3 паспорта специальности 03.01.09). По совокупности полученных результатов работа Храмеевой Екатерины Евгеньевны соответствует п.7 «Положения о порядке присуждения ученых степеней» ВАК МОН Российской Федерации, утвержденных Постановлением Правительства РФ № 842 от 24 сентября 2013 года. Результаты достоверны и достаточно полно опубликованы, автореферат адекватно отражает содержание диссертации. Автор заслуживает присуждения ученой степени кандидата биологических наук по специальности 03.01.09 - «Математическая биология, биоинформатика». Материалы диссертации были рассмотрены и обсуждены на Межлабораторном семинаре ИОГен РАН 17 октября 2014 года (протокол № 1).

5 ноября 2014 г.

Н.с. лаборатории системной биологии и вычислительной генетики
ИОГен РАН

к.ф.-м.н. Касьянов А.С.

Подпись научного сотрудника лабораторией системной биологии и вычислительной генетики ИОГен РАН, кандидата физико-математических наук Касьянова Артема Сергеевича удостоверяю

Ученый секретарь ИОГен РАН
д.б.н. Огаркова О.А.

