

## ОТЗЫВ

официального оппонента доктора биологических наук, заведующего лабораторией геносистематики Федерального государственного бюджетного научного учреждения науки «Лимнологический институт Сибирского отделения Российской академии наук» Щербакова Д. Ю. на диссертационную работу Бирюкова Михаила Юрьевича на тему «Происхождение и эволюция структурных вариантов *Tat* LTR-ретротранспозонов зелёных растений», представленной на соискание ученой степени кандидата биологических наук по специальности 1.5.7- «генетика»

**Актуальность темы диссертационного исследования** Диссертация Бирюкова М.Ю. посвящена исследованию эволюции мобильных элементов *Tat* из группы LTR ретротранспозонов, обнаруженных в геномах зелёных растений. Данная группа мобильных элементов имеет конвергентное сходство с ретровирусами позвоночных, дальними родственниками LTR ретротранспозонов, связанное с приобретением добавочного домена рибонуклеазы H в составе гена *pol*. Приобретённый домен в обеих группах выявлялся в аналогичном положении, следом за нативным доменом рибонуклеазы H, который, по-видимому, в какой-то момент утратил аминокислоты активного центра, выполняя у ретровирусов функцию домена-связки. У LTR ретротранспозонов *Tat* нативный домен рибонуклеазы H также частично терял аминокислоты активного центра, исходя из чего предполагается единая направленность и взаимосвязь процессов захвата добавочного и деградации нативного доменов. Однако, последовательность данных событий, их механизм и взаимосвязь остаются лишь предположениями. Группа LTR ретротранспозонов *Tat* была выявлена в геномах покрытосеменных растений, где все их копии, судя по положению добавочного домена рибонуклеазы H, имели сходную с ретровирусами структуру. Позднее были исследованы 3 генома хвойных голосеменных и один геном плауна, в которых были выявлены 2 структуры, отличающиеся по положению добавочного домена в *Tat*-LTR ретротранспозонах. Данные структуры рассматривались как исходные, по которым потенциально можно было бы проследить последовательность событий захвата добавочного домена и дальнейшей эволюционной судьбы домена нативного.

В данной диссертации автор производит углублённый поиск во всех основных таксонах зелёных растений и ближайших к ним таксонах водорослей, чтобы проследить момент возникновения и последовательность эволюционных событий, которые привели к возникновению ретровирусной структуры на LTR ретротранспозонах растений. Он также высказывает интригующее предположение о том, что полученные данные могут быть применены к исследованиям ретровирусов позвоночных, возможно, прошедших через аналогичные эволюционные преобразования.

**Научная новизна** Ряд важных результатов был получен автором впервые:

1. Для поиска, извлечения, аннотации и классификации LTR ретротранспозонов (в первую очередь с добавочным доменом рибонуклеазы H) был разработан автоматизированный алгоритм DARTS, основанный на поиске ключевых доменов LTR ретротранспозонов. В частности, на основе

поиска домена обратной транскриптазы алгоритм выявляет транспозоны групп Ty1/Copia, Ty3/Gypsy, Bel/Pao, DIRS, а также ретровирусы и Ренелоре-подобные элементы.

2. Впервые был произведён поиск *Tat* LTR ретротранспозонов среди геномов зелёных и стрептофитовых водорослей, печёночных и антоцетовых мхов, папоротников и древних таксонов семенных, таких как гингковые, саговниковые, гнетовые, кипарисовые, тисовые, амборелловые, нимфовые и магнолииды.
3. Было исследовано 94 генома зелёных растений и водорослей, среди которых было выявлено 5 структур по положению добавочного домена рибонуклеазы H, из которых 2 структуры описаны впервые.
4. Данными филогенетического и структурного анализа приведены свидетельства в пользу конвергентных процессов между *Tat* LTR ретротранспозонами, ретровирусами позвоночных и LTR ретротранспозонами оомицетов Chronos и Archon.
5. Приведены аргументы в пользу захвата добавочного домена рибонуклеазы H, как единичного события на *Tat* LTR ретротранспозонах, опровергая тем самым прежнюю гипотезу о множественном захвате из разных источников.
6. Был изучен процесс деградации нативного домена рибонуклеазы H при наличии в одном мобильном элементе одного или нескольких добавочных доменов с аналогичной функцией. На выборке кластеров внутри *Tat* показана связь между процессами захвата и деградации доменов. В частности, на примере структуры с тремя доменами, где два являлись добавочно приобретёнными, была также показана деградация одного из добавочных доменов, свидетельствуя о конкуренции между доменами, приводящей к деградации избыточных доменов даже между вероятно дублировавшимися копиями.

## Структура и объем диссертации

Диссертация имеет традиционную структуру: оглавление, список сокращений, введение, обзор литературы, материалы и методы, результаты и обсуждение, заключение, выводы и список литературы, который включает 157 источников. Работа изложена на 103 страницах, содержит 17 рисунков, 2 таблицы и 7 приложений.

**Введение** посвящено обоснованию актуальности проблемы, даны краткая характеристика мобильных элементов в целом и группы LTR ретротранспозонов *Tat* в частности, их сходства с ретровирусами позвоночных и сформулированы фундаментальные вопросы об их общих процессах эволюции на уровне «модулей» на примере белковых доменов. В данном разделе поставлена цель и определены задачи работы. Приведены основные характеристики работы, научная новизна, теоретическая и практическая значимость исследования, основные положения выносимые на защиту, а также личный вклад автора.

**Обзор литературы** содержит описание и общую характеристику мобильных элементов в широком смысле и отдельных групп LTR ретротранспозонов, содержащих добавочный домен рибонуклеазы H, в частности. Описаны жизненные циклы, особенности эволюции мобильных элементов и их влияние на эволюцию геномов организмов-хозяев. Подробнейшим образом рассмотрены белково-доменный состав *Tat* LTR ретротранспозонов и роль каждого из доменов в цикле воспроизведения мобильного элемента. Также подробно продемонстрировано актуальное состояние научных изысканий в ключевых вопросах исследования.

В разделе **Материалы и методы** подробно описаны биоинформатические методики, используемые в подобных исследованиях. На их основании автором разработаны, предложены и использованы новые подходы, сочетающие в себе как существовавшие прежде методики и программное обеспечение, так и их самостоятельную доработку, доведённую до автоматизированного алгоритма. В работе используется стандартный набор биоинформатического инструментария для проведения филогенетического анализа аминокислотных последовательностей, а также построения филогенетических деревьев. Выбор используемых методов исследования полностью соответствует поставленным в диссертации задачам. Полученные результаты прекрасно иллюстрированы, их достоверность не вызывает сомнений, а выводы полностью соответствуют полученным результатам.

Глава **Результаты** содержит описание сравнительного анализа эффективности разработанного автоматизированного алгоритма для поиска мобильных элементов, при сравнении его показателей с инструментами, использованными в прежних работах. С использованием данного алгоритма из 94 геномов зелёных растений было получено около 800 тысяч копий *Tat* LTR ретротранспозонов, содержащих добавочный домен рибонуклеазы H. На обнаруженных элементах были произведены филогенетический анализ по трём основным доменам и добавочному домену рибонуклеазы H, а также был произведён структурный анализ на предмет целостности, потенциальной активности элементов *Tat*, степени деградации нативного домена рибонуклеазы H. Был сделан вывод о преимущественно вертикальном наследовании данных элементов в эволюционной истории зелёных растений. Также была предпринята попытка выявить вручную пригодные для анализа битые и химерные *Tat* LTR ретротранспозоны и поиск добавочных открытых рамок считывания.

В главе **Обсуждение результатов** представлены результаты анализа филогенетических деревьев, построенных на основе доменов обратной транскриптазы, интегразы, нативного и добавочного доменов рибонуклеазы H, выявленных в целевых элементах. В первую очередь, разбиралось несоответствие филогении обратной транскриптазы и добавочного домена рибонуклеазы H, приводящих к заключению об единстве источника захвата и единственности данного события в эволюционной истории *Tat* LTR ретротранспозонов. Особое внимание было уделено анализу преобразования нативного домена в разных кластерах *Tat*, каждый из которых специфичен конкретным таксонам растений. По мнению автора, эти преобразования в целом носили характер деградации – утери исходной функции в результате мутаций в предполагаемом активном центре. В ходе обсуждения показана прямая взаимосвязь между захватом добавочного домена и исчезновением стабилизирующего отбора по последовательности нативного, поскольку в ходе работы также были выявлены кластеры элементов внутри *Tat*, одновре-

менно не содержащих добавочного домена, но сохранивших целым нативный домен рибонуклеазы H. Также проводится обсуждение обнаруженных вариантов расположения добавочного домена рибонуклеазы H, рассматриваются другие потенциальные позиции для добавочного домена и причины, почему, таковые не выявлялись или не были выгодны эволюционно. Глава завершается рассмотрением нескольких эволюционных сценариев, воссоздающих последовательность событий захвата, фиксации в различных позициях добавочного домена и деградацию нативного домена рибонуклеазы H в *Tat* LTR ретротранспозонах зелёных растений. Среди них выдвинут наиболее вероятный сценарий.

В **Заключении** подводятся итоги результатов работы и дискуссии вокруг полученных данных.

**Выводы** Выводы сформулированы ясно и полно, они соответствуют целям и задачам.

**Автореферат** . Текст автореферата соответствует содержанию диссертационной работы. По теме диссертации опубликовано 4 статьи (3 из которых - в журналах из перечня ВАК), а также тезисов сообщений по теме работы, которые были представлены на трёх международных конференциях.

## Замечания

1. Стр. 14 Диссертации. "От ретровирусов ведут своё происхождение протоонкогенные белки (вирусные киназы)". Избыточно категоричное утверждение. Следует писать "некоторые".
2. Стр. 33 и далее термин "гомология" используется в смысле "неслучайно сходные а общность происхождения предполагается по умолчанию, что приводит к возможной ошибке, особенно когда речь идет об относительно коротких последовательностях нуклеиновых или аминокислот.
3. Стр.36 При описании алгоритма поиска опущены критически важные детали. В частности, при описании второго построения дерева не указано, какую модель эволюции рассматривали - нуклеотидную, кодонную или аминокислотную, хотя очевидно, что они должны приводить к различным результатам. Разные значения должны быть и при использовании ультрабыстрого бутстрепа.
4. Стр.42 "...широкие возможности к адаптации DARTS... не указано, какие же параметры алгоритма следует менять
5. Сомнительно широкое использование термина "деградация". Требуется где-то в начале оговорить, что в данном контексте это - накопление отличий аминокислотной последовательности от канонической, которое вызывает (скорее всего!) ослабление или полную утерю исходной функции.

## Заключение

Таким образом, из вышесказанного можно сделать вывод о том, что диссертационная работа Михаила Юрьевича Бирюкова на тему «Происхождение и

