

ОТЗЫВ

официального оппонента на диссертационную работу

Бочкарёвой Ольги Олеговны

“Реконструкция эволюционной истории геномных перестроек в бактериях”,
представленную на соискание степени кандидата биологических наук по специальности
03.01.09 - “Математическая биология, биоинформатика”

Диссертационная работа Ольги Олеговны Бочкарёвой посвящена реконструкции и анализу геномных перестроек в различных бактериальных таксонах. Первые сравнения бактериальных хромосом были сделаны в начале 2000х годов для пар бактериальных хромосом с помощью точечных матриц. Эти работы показали, что порядок генов на бактериальных хромосомах слабо сохраняется на протяжении эволюции. Тогда же алгоритмические научные группы стали разрабатывать математические модели и алгоритмы для оценки эволюционных расстояний между штаммами в терминах перестроек.

К началу 2010 годов был разработан и опубликован ряд алгоритмов для восстановления и анализа геномных перестроек в группе штаммов, а также реконструкции филогении по данным о расположении порядка генов на хромосомах. В сочетании с доступностью в открытых базах данных десятков полностью собранных бактериальных геномов одного вида это позволило учитывать историю перестроек при анализе эволюции видов и родов, выявляя как общие закономерности, так и особенности определенных бактериальных таксонов. Тем не менее в микробиологических исследованиях работа с данными о порядке генов на хромосомах ограничивалась сравнением генома нового секвенированного штамма с геномом референсного вида и описанием выявленных различий, что показывает актуальность выбранной темы исследования.

Диссертация изложена на 111 страницах. Она включает девять разделов: введение, главы 1-5, заключение, список литературы и приложение. Глава 1 содержит обзор литературы по теме диссертации. Глава 2 содержит описание используемых методов и программ. Главы с 3 по 5 содержат описание собственных исследований. Работа содержит 37 рисунков и 5 таблиц. Список литературы содержит 112 наименований. Приложение содержит 5 таблиц.

Во **Введении** дается базовое представление об актуальности, научной новизне и практической значимости работы, сформулированы цели и задачи исследования, а также положения, выносимые на защиту.

Обзор литературы состоит из нескольких частей, полно описывающих современное состояние научной области. В первой главе кратко описана история развития методической базы и систематизированы возможности современных алгоритмов, использующихся для

работы с порядком генов на хромосомах. Во второй главе собраны описанные в литературе наблюдения и гипотезы о строении и эволюции бактериальных хромосом, необходимые для биологической интерпретации результатов биоинформатического анализа. Последняя часть обзора литературы посвящена описанию бактериальных таксонов, на геномах которых был выполнен анализ перестроек, с указанием особенностей их экологических ниш и строения геномов, что позволяет читателю понять, чем интересны те или иные виды с точки зрения эволюционной биологии и медицины. Значительная часть иллюстраций сделана автором диссертации самостоятельно (не заимствована из литературных источников).

Раздел **Материалы и Методы** содержит информацию обо всех алгоритмах и программах, использованных в исследовании, с указанием выбранных параметров запуска и биологических наблюдений, лежащих в основе использованных моделей. Раздел также содержит описание статистических моделей, использованных для оценки значимости результатов, полученных в диссертационной работе.

Результаты работы представлены в трех главах, первая из которых посвящена применению алгоритмов восстановления филогении штаммов по данным о расположении генов на хромосомах, вторая описывает результаты анализа паттерна перестроек для различных бактериальных таксонов, в третьей главе описана методика оценки фракции генов, распространяющихся в популяции горизонтально. Каждая из глав завершается обсуждением, в котором обсуждается значимость полученных результатов и возможное развитие этих направлений исследований. Полученные результаты хорошо обоснованы, рисунки и таблицы являются полезным для понимания сути работы иллюстративным материалом.

Результаты, полученные в диссертационной работе, вносят существенный вклад в понимание эволюции бактериальных хромосом, а подробно описанная автором методика может быть использована другими научными группами для изучения эволюции различных бактериальных таксонов. Работа охватывает широкий спектр эволюционных вопросов, таких как видообразование молодых патогенов, таких как *Yersinia pestis* и *Burkholderia mallei*, в которых именно перестройки генома сыграли ключевую роль; анализ действия эволюционного отбора по паттернам перестроек и выявление параллельных событий; оценка потока генов, горизонтально переносимых в популяцию, в штаммах, принадлежащим к разным экологическим нишам. Особой ценностью обладают результаты, полученные для мультихромосомных штаммов, поскольку структура и эволюция двукомпонентных бактериальных геномов остается слабоизученной и вопросы о механизме возникновения и стабильности таких геномов остаются открытыми. Описанный методический подход и наблюдения, сделанные на геномах *Burkholderia* spp., создают хороший задел для целого ряда

научных работ по изучению мультихромосомных геномов, которые в последние годы найдены в различных бактериальных таксонах.

Результаты, описанные в работе, представляют интерес не только с точки зрения исследований в области эволюционной биологии, но могут найти применение в практической микробиологии и медицине. В частности, описанный в работе подход может использоваться для биоинформатического поиска эволюционных событий, происходивших параллельно в ходе эволюции бактериального вида, и выявления инверсий, приводящих к антигенной вариации, и горизонтальных вставок геномных островков, приводящих к устойчивости к антибиотикам. Сделанное Ольгой Олеговной биоинформатическое предсказание о фазовой вариации за счет инверсии по паралограм *phtB* и *phtD* в штаммах стрептококков в дальнейшем было подтверждено в экспериментальной работе, опубликованной другими авторами, что показывает достоверность полученных результатов. Это особенно важно, поскольку в последние годы белок PhtD рассматривался как основной кандидат на роль мишени пневмококковой вакцины нового поколения.

К сожалению, автору не удалось избежать некоторого количества ошибок и опечаток. Ниже приводится список опечаток (выделены полужирным цветом) со страницами, на которых они присутствуют:

Стр. 4 - Коргеном (core-genome) — совокупность генов, присутствующих во всех организмов изучаемой группы.

Стр. 5 - Трансдукция (transduction) — процесс, при котором ДНК одной бактерии переносится в геном другой бактерии **засчет** встраивания фага.

Стр. 5 - Трансформация (transformation) — процесс поглощения клеткой свободной молекулы ДНК и **встаивания** ее в свой геном. — Тут следует отметить, что, как мне кажется, автор приводит не совсем правильную формулировку. Трансформация — это собственно процесс переноса в клетку ДНК из окружающей среды. Для этого клетка должна быть компетентной — это особое состояние, которое, в частности, можно вызвать искусственно. Дальнейшая судьба перенесенной в клетку ДНК может быть различной. В частности, при трансформации бактериальных клеток рекомбинантными плазмидами их встраивания в геном не происходит. Эта методика широко используется в генной инженерии.

Стр. 19 – В случае большого количество генов...

Стр. 23 - В случае же, если повторы лежат внутри генов, инверсия может привезти к замене части гена...

Стр. 24 - При этом сила отрицательного отбора на инверсии, лежащие внутри реплихоры, различается в разных видах бактерий и зависит от различных характеристик архитектуры

геномов, таких как **от** удаленность важных генов от точки начала репликации и степень перепредставленности генов на лидирующей цепи.

Стр. 31 – Виды *E. coli* **and** *Salmonella enterica* считаются близкородственными...

Стр. 39 – «в том время как в эволюции *Y. pestis* [59] происходят множественные геномные **геномные** перестройки (Таблица 3.1)».

Стр. 62 – «липопротеины внешней мембраны *nodT* **and** *emrA*».

Стр. 74 – «в истории которых **поисходил** горизонтальным перенос»; «переноса в **истори** гена».

К сожалению, автору не удалось избежать ошибок даже в выводе: Стр. 82 - Вывод 4 «в патогенах человека *Streptococcus pneumoniae* и *Burkholderia pseudomallei* предсказанны новые случаи антигенной вариации за счет инверсий».

Удивляет оформление списка литературы. Пример ссылки: *Kunin V., Ouzounis C. A.* The balance of driving forces during genome evolution in prokaryotes. // *Genome Res.* — 2003. — Т. 13, № 7. — С. 1589—94. В ссылках на публикации на английском языке (впрочем, других в диссертации нет) автор почему-то вместо V. (Volume) и P. (Pages) использует принятые для русскоязычных ссылок сокращения Т. (Том) и С. (Страницы).

При прочтении диссертации возник также ряд вопросов и замечаний, не связанных с опечатками в тексте.

Так, на стр. 20, Рис. 1-4 имеется ссылка на используемую программу <https://github.com/alex-bochkarev/dotplot.gen>. Эта программа была разработана в процессе выполнения работы и имеет отношение к диссертации? Ссылка не открывается. При рецензировании статей, содержащих ссылки на описываемые в статьях программы, рецензентов просят проверять, что ссылки открываются и программы работают. Здесь, вероятно, такой же случай?

На рис. 1-4 показана цепочка геномных перестроек. При этом нигде не написано, к каким именно геномам относится данная иллюстрация. Если приводится детальная иллюстрация, занимающая целую страницу, хотелось бы понимать, на каком материале она получена.

Стр. 27 - Рис. 1.6. – непонятно, что изображено на гистограмме. Что означают надписи – это названия штаммов или изолятов?

Термин «комменциал», употребленный на стр. 31 диссертации, встречается в Google всего 46 раз, причем в половине случаев – в статьях про велосипеды – это такой тип велосипедов. В русскоязычной литературе гораздо более распространенным является термин «комменсал» - встречается в Google 13400 раз, вероятно, его и следует употреблять.

Стр. 35 – Описан способ определения точки начала и конца репликации с помощью вычисления функции GC-skew – поскольку это не оригинальная методика, требуется привести ссылку.

Стр. 50 – Рис. 4.2 – Рисунок непонятен. Где находятся точки рекомбинации по результатам, приведенным на панели Б? В соответствии с приведенной в диссертации терминологией breakpoint – синоним точки разрыва, соответственно, в обсуждаемом месте – синоним точки рекомбинации. Каким образом рисунок может быть озаглавлен: «Поиск точки рекомбинации в брекпоинтах параллельной инверсии в стрептококках». Что в данном случае подразумевается под «брекпоинтами»? Почему в терминах написано по-английски, а в названии рисунка – по-русски?

Каким образом учитывались возможные ошибки сборки геномов при интерпретации результатов диссертации? Сколько подозрительных случаев, похожих на ошибочную сборку генома выявлено? Делались ли попытки исправления сборки – хотя бы для наиболее очевидных случаев? Может быть, об этом сообщалось авторам сборки?

Стр. 64 – Рис. 4.15. – В подписи написано «Цветными треугольниками помечены штаммы, в которых детектированы одинаковые инверсии». Однако, ни одного цветного треугольника на рисунке нет. Видимо, штаммы с одинаковыми инверсиями не детектируются. Тогда зачем было писать такую фразу в подписи к рисунку? Напротив, цветные треугольники присутствуют на Рис. 4.14, но в подписи к рисунку про треугольники ничего нет. Может быть, перепутаны подписи?

Ссылка на рис. 4.15 в тексте отсутствует. В тексте присутствуют фразы «Реконструкция выявила две параллельные инверсии на первых и одну — на вторых хромосомах. На границах этих инверсий также лежат геномные повторы, состоящие из генов, кодирующих транспозазы». Фразы, видимо, относятся к описанию геномов *B. thailandensis* и свидетельствуют в пользу того, что одинаковые инверсии в геномах есть.

При этом следует отметить, что все отмеченные замечания не носят принципиального характера.

Диссертационная работа Бочкарёвой Ольги Олеговны проведена на большом объеме данных, информация о которых указана в приложении к диссертации, и содержит значимые результаты, что подтверждается тремя статьями, опубликованными в международных рецензируемых научных журналах. Во всех трех статьях Ольга Олеговна является первым автором, что подтверждает, что работа выполнена автором самостоятельно, на высоком научном уровне с использованием современных биоинформатических методов и подходов. Содержание автореферата полностью отражает основные положения диссертации. Выводы диссертации обоснованны.

Таким образом, диссертационная работа Бочкарёвой Ольги Олеговны "Реконструкция эволюционной истории геномных перестроек в бактериях" является законченной научно-квалификационной работой и отвечает требованиям п.9 Постановления Правительства РФ «О порядке присуждения ученых степеней» №842 от 24 сентября 2013 г. с изменениями №335 от 21 апреля 2016 г., а ее автор заслуживает присуждения искомой степени кандидата биологических наук по специальности 03.01.09 - «математическая биология, биоинформатика».

Главный научный сотрудник лаборатории биологически активных наноструктур федерального государственного бюджетного учреждения "Национальный исследовательский центр эпидемиологии и микробиологии имени почетного академика Н.Ф. Гамалеи" Министерства здравоохранения Российской Федерации

доктор биологических наук,
профессор



А.С. Карягина-Жулина

20.04.2019

Москва, 123098, ул. Гамалеи, д. 18

Тел. +7 (499) 193-30-01, e-mail: akaryagina@gmail.com

Карягина-Жулина Анна Станиславовна

Подпись Карягиной-Жулиной А.С. удостоверяю

Ученый Секретарь Федерального государственного бюджетного учреждения "Федерального научно-исследовательского центра эпидемиологии и микробиологии имени почетного академика Н.Ф. Гамалеи" Министерства здравоохранения Российской Федерации
кандидат биологических наук



Л.К. Кожевникова