

Отзыв официального оппонента о диссертационной работе
Манолова Александра Ивановича
«Биоинформатический анализ изменчивости генного состава прокариот, в том числе в ассоциации с патогенностью»,
представленной на соискание ученой степени кандидата биологических наук по специальности 1.5.8. – математическая биология, биоинформатика.

Изучение структур в геномных последовательностях имеет важное значение для понимания биологических процессов, происходящих в живых клетках. Объединение генов в опероны, неравномерное распределение расположения генов по двум цепям ДНК – примеры уже давно известных закономерностей организации генетического материала. Представленная к защите диссертационная работа посвящена исследованию менее изученного вопроса о локализации изменений, происходящих в результате горизонтального переноса генов и геномных перестроек, в геномной последовательности. Одним из объектов исследования в данной работе являются области повышенной изменчивости, в которых наблюдаемое количество замен одних генов на другие значительно превышает фоновое значение. Для выявления подобных областей был разработан метод, основанный на применении графов. Данный метод позволил оценить уровень изменчивости в различных областях геномной последовательности. Важной частью исследования стало сравнение профилей изменчивости различных организмов, которое показало консервативное расположение ряда высоко-изменчивых областей. Помимо численной оценки уровня изменчивости, в работе решается задача визуализации сочетаний генов внутри геномных последовательностей в виде графа, что позволяет сравнивать расположение генов в сотнях геномных последовательностях. Предложенный способ позволил выявить характер изменений, наблюдающихся в областях с повышенной и пониженной изменчивостью. Было показано существование фрактало-подобных структур, когда внутри высоко-изменчивых областей встречаются, вложенные в них, высоко-изменчивые области, окруженные консервативным контекстом.

Предложенные новые методы исследования геномных последовательностей были применены для изучения изменчивости штаммов кишечной палочки, ассоциированных с развитием воспалительного заболевания кишечника - болезни Крона. В работе дается характеристика изменчивости в областях оперонов, которые значимо чаще встречались в изолятах кишечной палочки, полученных от пациентов, по сравнению с изолятами от здоровых людей. Показано, что ряд подобных областей обладает высоким уровнем изменчивости, в частности – области локализации оперонов утилизации пропандиола, а также синтеза и транспорта капсулы.

Работа изложена на 131 странице и состоит из шести глав, включает в себя 45 рисунков и две таблицы, библиография содержит 231 наименование. По материалам диссертации опубликовано шесть статей, результаты работы представлены на шести конференциях.

Первая глава посвящена обзору литературы, в котором автор рассматривает основные известные закономерности организации генетического материала, наиболее изученные механизмы горизонтального переноса генов, биоинформатические подходы к изучению геномных последовательностей и связь между болезнью Крона и бактерией *Escherichia coli*. Во второй главе указаны методы и программные продукты, использованные для сбора геномной информации.

В третьей главе приведены результаты работы. Вначале излагаются общие принципы представления геномной информации в виде графа. Предложено представлять группы ортологичных генов в качестве узлов графа и использовать информацию о смежном расположении генов в одной из геномных последовательностей для связи соответствующих узлов графа. Следующим шагом предлагается использовать построенный таким образом граф для оценки уровня изменчивости генома. Оценка производится на основе подсчета различных разветвлений в графе, которые отражают смену генного состава вследствие вставок, удаления либо транслокации генов. Предложенный подход был проверен на данных, полученных компьютерной

симуляцией изменений в геномных последовательностях, а также сравнением с ранее опубликованными результатами. Далее приводится информация о чувствительности получаемых оценок к количеству и качеству используемых геномных последовательностей. Третий пункт главы посвящен результатам применения разработанного подхода оценки уровня вариабельности. На основе рассмотрения более чем трехсот геномов кишечной палочки показано, что большая часть областей повышенной изменчивости ассоциирована с мобильными элементами генома и профагами. Отмечено существование как минимум одной области с высоким уровнем изменчивости, но не содержащей опознаваемых признаков мобильных элементов генома. Были проанализированы профили изменчивости геномов из различных клад филогенетического дерева *E. coli* и показано наличие устойчивых во времени областей повышенной изменчивости. Подобная методология - сравнение профилей устойчивости с применением информации о сходстве нуклеотидных последовательностей - была также применена для сравнения внутривидовых структур у видов *Pseudomonas fluorescens* и *Pseudomonas aeruginosa*. Данное сравнение показало низкий уровень сходства профилей изменчивости в первом случае, и высокий - во втором. Обнаруженное различие объясняется в работе наличием высокого уровня крупномасштабных геномных перестроек у *P. fluorescens*, что по мнению автора значительно ограничивает применимость предложенного метода анализа локальной изменчивости. Также приводится сравнение профилей изменчивости для различных видов внутри одного рода. Такие сравнения проведены для четырех видов рода *Bacillus*, и для пар видов внутри родов *Chlamidia* и *Lactobacillus*. Эти сравнения показали, что консервативное расположение участков повышенной изменчивости может наблюдаться и при сравнении различных видов, и подтвердило значительный вклад профагов в формирование ландшафта изменчивости геномной последовательности. В четвертой части третьей главы предпринята попытка найти факторы, ассоциированные с уровнем изменчивости на примере генома кишечной палочки. Сравнение расположения сайтов Chi, играющих важную роль в процессе гомологичной рекомбинации, показало наличие статистически значимой ассоциации с уровнем изменчивости: чем

реже расположены данные сайты, тем, в среднем, выше уровень изменчивости. К сожалению, остается неясным является ли концентрация этих сайтов причиной, вызывающей повышение уровня изменчивости, или наблюдаемая связь является следствием более редкой встречаемости последовательностей Chi в мобильных элементах генома. Для бактерий *E. coli* и *Bacillus subtilis* было проведено сравнение уровня изменчивости и плотности межхромосомных контактов, полученными из данных литературы. Наблюдалась статистически значимая, но при этом низкая по абсолютному значению коэффициента, корреляция. Следующая часть главы посвящена подходу к визуализации в виде графа расположения генов в геномных последовательностях. Предложенный метод позволяет оценить зрительно уровень геномной изменчивости в выбранном локусе и характер этих изменений. Приведены визуализации для областей генома кишечной палочки в которых расположены опероны утилизации пропандиола, синтеза и транспорта капсулы, захвата гемина, утилизации глиоксилата и сорбозы. В седьмой части автор представляет разработанное в ходе выполнения работы приложение - Genome Complexity Browser, позволяющее одновременно иметь доступ и к профилю изменчивости всей геномной последовательности и построению графов для выбранных областей генома. В финальной части главы, автор проясняет метод отбора локусов генома кишечной палочки, ассоциированных с наличием болезни Крона у пациентов, являющихся носителями этих бактерий.

Четвертая глава диссертации посвящена обсуждению работы и содержит, помимо перечня основных результатов, детальное описание недостатков и ограничений предложенной методологии. К недостаткам метода автор причисляет сложность работы с паралогичными генами, к ограничениям – требование к сохранности крупномасштабной архитектуры геномных последовательностей для возможности анализа локальных изменений.

Ниже приведено несколько замечаний и вопросов:

1. Одной из главных проблем в работах, важной целью которых является поиск связей между сложными феноменами, такими, как например, в данной работе поиск связи между этиологией болезни Крона и особенностями геномов *E.*

coli, является комплексная взаимообусловленность факторов. Так, например, комплекс специфических реакций/метаболитов пациента, характерный для болезни Крона, может стать фактором, благоприятствующим развитию специфических генотипов *E.coli* (что отчасти довольно правдоподобно, если учесть данные о «перепредставленных» оперонах из вывода 4). Если это так, то следовало бы уточнить к какой более крупной цели, чем просто индикация особенностей штаммов *E.coli*, выделенных из пациентов с болезнью Крона, направляется данное исследование;

2. Аналогичная проблема связана с распределением хот-спотов в геноме бактерий. Известно (и Автор отмечает это в литературном обзоре), что интеграция геномных островов часто происходит в гены транспортной РНК. Геномный остров, характеризующийся, как правило, высоким содержанием IS-элементов, становится после этого рекомбинационным хот-спотом. Однако причиной локализации этого хот-спота является наличие лишь гена тРНК, а не какой-либо другой особенности данного локуса. В связи с этой проблемой между прочим возникает также вопрос о том, что Автор не выявил явной ассоциации генов транспортных РНК с профилем изменчивости. Верно ли, что геномные острова в проанализированных геномах *E.coli* интегрированы по альтернативным сайтам?
3. Уточните, каким образом были идентифицированы границы реплихор в анализируемых геномах и куда делись пунктирные линии, обозначающие их границы на рис. 3.18 (возможно, я их просто не заметил)?
4. В качестве контрольных видов микроорганизмов, геномы которых были проанализированы Автором, были помимо прочих использованы геномы *N. gonorrhoeae*, общепризнанной особенностью биологии которых является большая склонность к естественной трансформации и, как следствие, чрезвычайно высоким уровнем внутрипопуляционной рекомбинации. Каким образом эта особенность данной группы микроорганизмов может быть связана с особенностями профиля изменчивости в данной группе?

5. Наконец, вопрос, связанный с выбором референса. Авторами в качестве референса был выбран геном штамма, выделенного из пациента с болезнью Крона. Каким образом выбор альтернативного референса может (и может ли) отразиться на структуре графа и мерах локальной изменчивости (числе путей и весов ребер)?

Все вопросы, сформулированные в данной части отзыва не имеют критического характера и не вызваны сомнениями в ценности этой очень интересной и хорошо сделанной работы, их цель лишь уточнение некоторых положений.

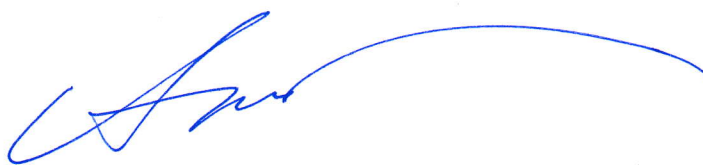
Заключение. Диссертационная работа Манолова А.И. посвящена актуальной теме - исследованию закономерностей процессов изменений генного состава в геномах прокариот, в частности исследованию областей с повышенным уровнем изменчивости. Исследование стало возможным благодаря предложенным автором оригинальным методам численной оценки уровня изменчивости и визуализации изменений в наборе генов. Выполнен анализ изменчивости геномных последовательностей у бактерий вида *E. coli* и ряда других микроорганизмов. По теме диссертации опубликовано шесть публикаций. Все это позволяет рекомендовать автора работы А.И. Манолов для присуждения степени кандидата биологических наук по специальности 1.5.8 – математическая биология, биоинформатика. Работа полностью удовлетворяет требованиям к кандидатским диссертациям п. 9 “Положения о порядке присуждения ученых степеней”, утвержденного Постановлением Правительства Российской Федерации от 24 сентября 2013 г. N 842 "О порядке присуждения ученых степеней" (с изменениями в редакции Постановления Правительства Российской Федерации от 21 апреля, 2 августа 2016 г., 29 мая, 28 августа 2017 г., 1 октября 2018 г., 20 марта, 11 сентября 2021 г.).

Официальный оппонент:

Заведующий лабораторией микробиологического мониторинга и биоремедиации почв ФГБНУ «Всероссийский научно-исследовательский институт сельскохозяйственной микробиологии» (г. Санкт-Петербург), кандидат биологических наук, специальность 03.00.07– «Микробиология». 196608, Санкт-Петербург, Пушкин, ш. Подбельского, 3, эл. почта eeandr@gmail.com тел. +7-911-213-93-01

Андронов Евгений Евгеньевич

19 мая 2022 года



Подпись руки Андропова Евгения Евгеньевича, ведущего научного сотрудника лаборатории микробиологического мониторинга и биоремедиации почв ФГБНУ ВНИИСХМ, кандидата биологических наук, заверяю: начальник отдела кадров ФГБНУ ВНИИСХМ Ковалевская М.А. 19 мая 2022 года.

