

УТВЕРЖДАЮ

Генеральный директор

ФБУН ГНЦ ВБ «Вектор»

Роспотребнадзора,

доктор биологических наук

Р.А. Максютлов

« 24 » сентября 2021 г.



ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

на диссертационную работу

Мустафина Захара Сергеевича

на тему «Разработка комплекса программ для анализа эволюционных характеристик генных сетей», представленной к защите в диссертационном совете 24.1.239.01 на базе ФГБНУ «Федеральный исследовательский центр

Институт цитологии и генетики СО РАН» на соискание ученой степени кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика

Актуальность избранной темы, теоретическая и практическая значимость

Диссертационная работа Мустафина Захара Сергеевича посвящена разработке программного обеспечения для анализа эволюционных характеристик генных сетей – приложений Orthoscape и Orthoweb. Orthoscape является плагином к Cytoscape и предназначено для анализа эволюционных характеристик генов в генных сетях, а именно: (1) анализ с целью выявления, являются ли гены гомологичными; (2) поиск предполагаемого этапа возникновения гена на таксономическом дереве; (3) определение уровня

эволюционной изменчивости гена. Orthoweb – веб-приложение со схожей с Orthoscape функциональностью, ориентированное на анализ наборов генов, не объединенных в генную сеть. В настоящее время наблюдается дефицит программного обеспечения для эволюционного анализа генных сетей, таким образом, актуальность представленной работы не вызывает сомнений.

Анализ эволюционных характеристик генов, входящих в различные генные сети, регулирующих различные процессы в организме, позволяет построить картину их развития, понять, какому типу отбора подвергались гены, входящие в генную сеть, определить, какие гены в тех или иных процессах являются наиболее древними/молодыми и в то же время, являются ли эти гены консервативными или же, наоборот, изменчивыми, позволяет выделить группы генов, интересных для более подробного анализа.

Новизна исследования и полученных результатов, выводов и рекомендаций, сформулированных в диссертации

Метод филостратиграфического анализа был впервые применен для анализа генных сетей. Впервые разработаны и реализованы программы для филостратиграфического анализа генных сетей, и проведен филостратиграфический анализ генных сетей заболеваний человека различной природы и анализ генных сетей, регулирующих ответ на стресс у *A. thaliana*.

Структура диссертационной работы

Диссертация изложена на 116 страницах печатного текста и состоит из введения, литературного обзора и заключения к нему, описания материалов и методов, двух глав с описанием результатов моделирования, заключения, выводов и списка литературы; содержит 6 таблиц и 42 рисунка. Список цитированной литературы содержит 139 источников. Структура диссертации построена по традиционному плану. Текст состоит из введения, пяти глав, заключения, выводов и списка литературы.

Во введении автор убедительно обосновывает актуальность выбранного направления исследований. Цели и задачи исследования, положения,

выносимые на защиту, сформулированы корректно и соответствуют теме диссертационной работы.

В первой главе, посвященной обзору литературы, излагается современное состояние исследований в области филогенетического анализа и применимости данного подхода для анализа генных сетей. Приведено описание ряда основополагающих баз данных и знаний. Приводится обзор ряда эволюционных индексов. Тем не менее, несмотря на обширность изложенного материала, обзор литературы оставляет впечатление незаконченного и несколько поверхностного.

Во второй главе, посвященной материалам и методам, описаны использованные алгоритмы, описаны эволюционные индексы, достаточно подробно описаны технологии, использованные при разработке приложений.

В третьей главе подробно описываются разработанные автором программы для анализа генных сетей – Orthoscape и Orthoweb, предоставляющие пользователю возможность анализа генных сетей или самостоятельных наборов генов с целью установления между генами отношений гомологичности, осуществления филогенетического анализа и определения типа отбора, которому подвержены гены в процессе эволюции организма. Описаны их функциональные возможности, параметры запуска. Приведены примеры филогенетического анализа с помощью разработанных эволюционных индексов. Описаны входные данные и структура выходных данных.

В четвертой главе описаны результаты анализа генных сетей заболеваний человека, показавший, что в сетях группы «Immune diseases» в большинстве своем задействованы «молодые» гены, обладающие высоким значением индекса дивергенции при сравнении с генами группы гоминид, что говорит о том, что эти гены активно эволюционируют. В ходе анализа генных сетей, вовлеченных в формирование зависимостей от различных веществ («Substance dependence») показано, что большинство генов, входящих в данные сети, являются эволюционно «древними», что говорит о

консервативности и эволюционной древности процессов, лежащих в основе этой группы заболеваний.

В пятой главе представлены результаты анализа генов, ассоциированных с различными типами стресса у *A. thaliana*. Автором с помощью Orthoscape и Orthoweb были проанализированы списки генов, которые ассоциированы со стрессом у *A. thaliana*. Из базы данных TAIR v. 10 были извлечены списки генов, ассоциированных с холодным, тепловым, световым, осмотическим, солевым, водным и оксидативным стрессом; на основе составленных списков были построены генные сети, которые были проанализированы. Показано, что древние гены достоверно перепредставлены, а молодые недопредставлены в выборках генов, ассоциированных со стрессом у *A. thaliana*. Полученные результаты также свидетельствуют о многофункциональности эволюционно древних генов, участвующих в реакции на стресс.

В заключении подводятся итоги выполненного исследования и описываются основные результаты. Достоверность полученных автором выводов не вызывает сомнений. Основные результаты были опубликованы в трех статьях, в рецензируемых научных журналах, входящих в список ВАК, а также были представлены на международных научных конференциях. Получено авторское свидетельство. Список литературы включает 139 ссылок.

Замечания

В работе имеется ряд недостатков. В частности, присутствуют неточные формулировки, например на стр. 5 присутствует фраза «Целью филостратиграфического анализа является определение времени возникновения гена на основе оценки распределения ортологичных ему генов в геномах организмов, принадлежащих к различным таксономическим группам», в то время как в данной работе время оценивается косвенно. Еще один пример: «к настоящему моменту не существует общепризнанного метода расчёта филостратиграфического индекса, имеющего стабильную программную реализацию» (стр. 5, 24). На стр. 14 приведено соотношение

dN/dS, а пояснение приводится лишь на стр. 21. Иногда встречаются английские слова, использование которых не представляется необходимым: на стр. 15 слово «score» можно заменить на «значение» или другой подходящий синоним. На стр. 16-17 упоминается ряд проблем при использовании BLAST для поиска гомологов, на мой взгляд, здесь нужно было бы добавить больше конкретики. Указано, что некорректные результаты отмечались различными коллективами авторов, но приведено лишь 2 ссылки – без обсуждения результатов, описанных в этих статьях, и без указания версий программного обеспечения, параметров запуска.

На стр. 19 указано, что «Не существует стандартного филогенетического дерева для осуществления филогенетического анализа, независимо от организма. Исследователь должен самостоятельно определять степень точности выделения таксонов в зависимости от имеющихся у него данных и прочих факторов.» Это существенное ограничение при сопоставлении результатов – хорошо было бы отметить и обсудить этот факт.

В главе 2 (стр. 40) не указана версия лицензии GPL, некорректно оформлены ссылки на ресурсы сети интернет, для ряда ресурсов и программ не приведены литературные ссылки.

Ряд рисунков имеет низкое разрешение и плохо читается. Попадают опечатки и пунктуационные ошибки, правда, очень немногочисленные.

У пятой главы, в отличие от предыдущих четырех, отсутствует заключение.

Однако все указанные выше замечания носят рекомендательный характер и несколько не умаляют научной значимости представляемой работы и полученных результатов.

ЗАКЛЮЧЕНИЕ

Тема диссертационной работы Мустафина З.С. «Разработка комплекса программ для анализа эволюционных характеристик генных сетей» соответствует научной специальности 03.01.09 (Математическая биология, биоинформатика). Работа посвящена актуальной задаче – разработке программного обеспечения для филостратиграфического анализа генных сетей, представляющей значительный интерес с точки зрения анализа эволюции регуляции различных биологических процессов, в том числе, лежащих в основе различных заболеваний и т.д. Научная новизна положений, выносимых на защиту, не вызывает сомнений. Полученные в ходе выполнения диссертационной работы научные и практические результаты полностью соответствуют поставленным целям. Содержание диссертации должным образом отражено в автореферате и в опубликованных работах. Выводы диссертации адекватно отражают полученные результаты.

Представляемая работа была апробирована на многих международных конференциях, получено авторское свидетельство, материалы диссертации опубликованы в трех статьях в рецензируемых научных журналах, входящих в перечень ВАК.

Захар Сергеевич продемонстрировал способность к глубокому системному анализу наряду с высокой квалификацией в области информационных технологий. Текст диссертации написан грамотным научным языком.

Таким образом, работа Мустафина З.С. «Разработка комплекса программ для анализа эволюционных характеристик генных сетей» полностью соответствует требованиям п.9 «Положения о порядке присуждения ученых степеней», утвержденного Постановлением Правительства РФ от 24.09.2013 г. №842, предъявляемым к диссертациям на соискание ученой степени кандидата наук, а ее автор заслуживает присуждения искомой ученой степени – кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика.

Отзыв на диссертацию и автореферат Мустафина Захара Сергеевича
обсужден и принят на заседании Ученого совета Федерального Бюджетного
учреждения науки «Государственный научный центр вирусологии и
биотехнологии «Вектор» Федеральной службы по надзору в сфере защиты
прав потребителей и благополучия человека» (Протокол № 7 от 24.09.2021 г.).

Кандидат биологических наук,
старший научный сотрудник
теоретического отдела,

ФБУН ГНЦ ВБ «Вектор» Роспотребнадзора



Д.В. Антонец

Подпись Д.В. Антонца заверяю

Начальник отдела кадров

ФБУН ГНЦ ВБ «Вектор» Роспотребнадзора



И.В. Ильин

Федеральное бюджетное учреждение науки «Государственный научный центр
вирусологии и биотехнологии «Вектор» Федеральной службы по надзору в
сфере защиты прав потребителей и благополучия человека»

630559, Новосибирская область, р.п. Кольцово,

+7(383)336-60-10

antonec@yandex.ru