

Отзыв официального оппонента

на диссертационную работу

Романова Дмитрия Евгеньевича

«Анализ связи между структурно-функциональной организацией генома в окрестностях регулирующих рост генов и морфо-физиологическими характеристиками млекопитающих»,
представленную на соискание ученой степени кандидата биологических наук по специальности «03.02.07 – генетика»

Диссертационная работа Д.Е.Романова представляет оригинальное исследование по фундаментальной проблеме генетике и биологии развития - оценке структурных и молекулярно-генетических механизмов, определяющих рост тела и морфо-физиологические характеристики с использованием методов биоинформатики.

Продолжающееся развитие высокопроизводительных геномных технологий дает новые данные в модельных организмах млекопитающих для анализа генетических механизмов, определяющих такие морфо-физиологические характеристики, как массу и размер тела взрослого животного, период полового созревания и продолжительность жизни. Необходима разработка биоинформационных методов и программ систематизации таких данных для конкретных задач, успешно выполненных соискателем в данной работе.

Актуальность исследования

Генетическая регуляция роста и размера тела остается неисследованной проблемой генетики. Известно, что скорость соматического роста млекопитающих высока в ранние этапы развития организма, но с возрастом постепенно снижается, задавая тем самым конечный размер тела взрослого животного. Снижение скорости роста млекопитающих, заключающееся в уменьшении уровня пролиферации клеток в разных тканях, контролируется в основном локальными механизмами регуляции, исследование которых важно для биомедицинских проблем. Результаты данной работы востребованы в исследованиях генетических нарушений, вызывающих задержку роста или преждевременное старение. Таким образом, исследование актуально.

Структура и основное содержание диссертации. Диссертация построена традиционным образом и состоит из следующих разделов: введения, обзора литературы, описания материалов и методов исследования, изложения результатов и их обсуждения, заключения, выводов, списка сокращений и списка цитируемой литературы (145 источников). Работа представлена на 123 страницах и содержит 11 рисунков и 36 таблиц.

Работа состоит из двух логически разделенных частей. Первая часть посвящена разработке ряда программных средств для работы с геномными данными. Предложен компьютерный метод поиска консервативных элементов генома, позволяющий проводить множественное сравнение геномных последовательностей.

Во второй части с помощью разработанных автором программных средств был выполнен анализ окрестностей семнадцати генов, регулирующих рост и размеры тела млекопитающих.

Степень обоснованности и достоверность научных положений и выводов. Работа выполнена на высоком методическом и техническом уровне. Используются актуальные версии баз данных и репозиториев, отражающих имеющиеся на данный момент знания о генах роста и консервативных элементах генома. Тем самым проанализирована большая часть имеющейся геномной информации. Выводы работы полностью обоснованы и логически вытекают из полученных результатов.

По материалам диссертационного исследования опубликовано 18 печатных работ: 3 статьи в изданиях Scopus, 2 свидетельства о регистрации программ для ЭВМ, 13 тезисов в сборниках трудов конференций

Новизна исследований и полученных результатов.

Диссертантом разработана модификация компьютерного метода поиска консервативных элементов генома на основе построения и анализа точечной матрицы гомологий, позволяющая проводить множественное сравнение геномных последовательностей. Выявлены консервативные участки генома в окрестностях регулирующих рост генов у различных видов млекопитающих.

Впервые показана связь между геномным расстоянием между консервативными элементами в окрестностях регулирующих рост генов и такими морфо-физиологическими характеристиками млекопитающих, как масса и длина тела животного и продолжительность жизни.

Показано, что геномное расстояние между регуляторными элементами генома млекопитающих может выступать фактором, эволюционно модулирующим экспрессию генов регуляции роста.

Впервые выявлена взаимосвязь между морфо-физиологическими характеристиками млекопитающих и положением на хромосоме регулирующих рост генов.

Теоретическая и практическая значимость полученных результатов

По результатам исследования обнаружена зависимость между морфо-физиологическими характеристиками животных и геномным расстоянием между некоторыми консервативными элементами в окрестности регулирующих рост генов *Mycn*, *Plagl1* и *Ezh2*.

Показано наличие значимой корреляции между периодом полового созревания и геномным расстоянием от начала гена до теломеры для генов *Ghrh*, *Sst*, *C1s* и *Notch1*. Предложены молекулярные механизмы, объясняющие обнаруженные зависимости.

Согласно выводам данного исследования, геномное расстояние между регуляторными элементами генома может являться фактором, эволюционно модулирующим экспрессию генов регуляции роста и определяющим фенотипические различия между различными видами млекопитающих.

Результаты работы могут быть востребованы для селекции с целью повышения морфо-физиологических показателей сельскохозяйственных животных. Результаты значимы для медицины, для анализа генетических нарушений, связанных с нарушениями роста организма.

Замечания и рекомендации.

Среди достоинств работы необходимо отметить важность проблемы, новизну подхода, тщательную методологическую проработку задачи. Однако, представляется целесообразным высказать ряд пожеланий и замечаний к работе:

Автореферат, содержание которого соответствует диссертации, имеет увеличенный объем - 25 страниц.

Вопрос по статистическому обоснованию - «На основании визуального анализа гомологий были выявлены консервативные элементы». Визуально - это не обоснование. Должна быть статистическая оценка такого выделения. Подразумевается, что такая оценка уже заложена в программе, и визуализация лишь показывает результат. Этот момент нужно прописать более четко.

Фраза «Поиск... известных элементов генома (ИЭГ)». Должно быть четкое определение, что значит «известных»? То есть известных на сегодняшний день из таких-то баз данных. Вводится новое аббревиатура, определение, без ссылки. Формулировки должны быть четкими, особенно в

выводах работы - известных автору, или известных из научной литературы. Представлена таблица используемых геномных баз данных, но данные не сгруппированы по классам - это элементы генома, относящиеся к регуляторным районам (сайты связывания транскрипционных факторов, определенные различными методами, псевдогены, участки нуклеотидных полиморфизмов генома - все они имеют разную природу).

Реализован ряд скриптов для автоматизированного извлечения информации из баз данных NCBI. Однако, описание собственных программных средств представлено в краткой форме. Кроме того, исходный код некоторых разработанных программных средств не выложен в открытый доступ.

Детали использования стандартных программ не указаны, например BLASTN ; чувствительность поиска: Distant homologies). Какая версия программы BLAST использовалась - NCBI, или EMBL? Стоит указать версию, параметры поиска (предполагается, что был поиск удаленных гомологов со слабыми ограничениями).

«Эпигенетические часы Хорвата» - термин используется в тексте, но само пояснение дано позже. Следует хотя бы назвать по-другому - «Эпигенетические часы, предложенные в работах Стивена Хорвата (Horvath S., 2013)». Да, это известная работа для специалистов, но фамилия автора без ссылки может вводить в заблуждение.

Также, обсуждаемый «теломерный эффект положения на длинных расстояниях (telomere position effect over long distances, TPE-OLD)» в тексте пояснен со ссылками позже, чем первое упоминание (стр. 3 автореферата). Стоит пояснить эффект хотя бы одним предложением сразу, при первом упоминании. Иначе непонятно, почему использовался именно этот список генов. То есть все обозначения и термины должны поясняться сразу. Ссылки на упоминаемые работы [25,25] идут на один и тот же номер и том журнала Genes and Development за 2014 год, который записан по-разному. Более свежие работы (2016г.) процитированы в диссертации, но стоило бы процитировать и развитие подхода, например, статьи последних лет (Kim and Shay. Long-range telomere regulation of gene expression: Telomere looping and telomere position effect over long distances (TPE-OLD). Differentiation. 2018).

В выводах 5 и 6 представлены избыточные детали и описание категорий генных онтологий на английском, вместо формулировки по существу. «Анализ сверхпредставленности категорий Gene Ontology по этим генам выявил гены ACVR1B и ACVRL1, значимо представленные в категориях growth factor binding, GO:0019838 (p = 0.0988) и response to growth factor,

GO:0070848 ($p = 0.0461$)). Корректнее сформулировать как: «Анализ сверхпредставленности категорий генных онтологий по этим генам выявил гены ACVR1B и ACVRL1, значимо представленные в категориях «связывание ростовых факторов» и «ответ на ростовые факторы»», возможно с оригинальными терминами на английском, но точных без цифр статистики (которые и затрудняют чтение, и зависят от использовавшейся методики подсчета).

В таблице 5 (автореферат) названия колонок не корректны - «Клин. Знач dbSNP». Это не клиническое значение, а идентификатор базы данных dbSNP. То же замечание - для ClinVar.

В таблицах 9 и 10 (автореферат) (Корреляция между морфо-физиологическими характеристиками млекопитающих и абсолютным / относительным геномным расстоянием...) статистические значения 0.01 выглядят неубедительно даже после коррекции на множественность гипотез.

Нет пояснения что такое ($x 200$). В тексте диссертации приведено подробнее, но также без пояснений по таблицам.

Общее замечание по оформлению самого текста диссертации - Таблицы занимают большой размер в тексте, по три страницы подряд занято таблицами без подписей и с разрывом текста, что очень затрудняет чтение.

Есть другие стилистические погрешности -

Например - «механизмы, ... остаются открытым вопросом» - правильнее было бы сказать - «Исследование механизмов... остается открытой проблемой»

«Работа является шагом на пути к пониманию, как связаны...» - наверное, лучше сказать «Работа является шагом на пути к исследованию того, как связаны...»

В разделе «Личный вклад» написано:

«По мере выполнения работы был реализован ряд методов...», но не написано кем. Наверное, надо отметить, кем реализованы эти методы - лично автором, или в коллективе?

Есть и опечатки. В автореферате и диссертации - Gene Ontology - должно быть gene ontology что впрочем, понятно из контекста.

Опечатки (падежи) в положениях, выносимых на защиту - «корреляция между морфо-физиологическими характеристиками» Надо «морфо-физиологическими».

Задача 6. «зависимость между ... и положением.. регулируемыми рост генов» - надо «.регулирующих рост генов»

Фраза «Показано, что уменьшение с возрастом релизинга GH1 связано с уменьшением релизинга GHRH и увеличению релизинга SST...» - наверное, с

«увеличением релизинга SST...» - и лучше сказать «статистически связано», обосновав такой вывод.

Список сокращений дан слишком кратко - некоторые обозначения по тексту не показаны в списке обозначений и сокращений.

Формулировка по списку литературы (в автореферате) - заголовок «Научные статьи по теме диссертации, опубликованные в журналах Scopus, WoS, RSCI» - корректнее сказать «издания», не «статьи». Издание [3] не является статьей в журнале Abstracts from the 51st European Society of Human Genetics Conference: Electronic Posters. Этот момент влияет и на общие параметры представления диссертации - то есть автором опубликовано 2 статьи в журналах, индексируемых в Scopus, и 3 публикации в изданиях индексируемых в Scopus.

Замечу досадную, но принципиальную опечатку в списке литературы по тезисам Школы молодых ученых и конференции серии BGRS в Новосибирске «Под ред. А. В. Харкевич.—г. Новосибирск : ФГУП «Издательство СОРАН»». Как представитель и член оргкомитета данной серии конференций BGRS с 2000 года, отмечу что все труды издавались под редакцией академика Н.А.Колчанова, либо коллективно оргкомитета, а А.В.Харкевич - это инженер, выполнявший компьютерную верстку, не научный редактор, что и указано в выходных данных трудов и сборников. Таким образом, выходные параметры публикаций четырех изданий указаны не верно.

Отметим, что есть другие публикации соискателя в журналах, которые, на мой взгляд, относятся к теме работы, но не отмечены в списке литературы именно по теме диссертации. Таким образом, по существу, все научные результаты опубликованы, в соответствии с требованиями к кандидатским диссертациям.

Сделанные замечания не снижают общей высокой оценки диссертационной работы.

Общее заключение.

Содержание диссертации полностью отражено в публикациях автора в журналах, рекомендованных ВАК, а также в других рецензируемых изданиях, и представлено на международных и российских научных конференциях. Основные результаты и их интерпретация представлены в автореферате, который адекватно отражает содержание диссертации. Поставленные в диссертации задачи решены полностью.

Таким образом, соискатель, Романов Дмитрий Евгеньевич, обладает необходимыми компетенциями для формулирования и решения

исследовательских и практических задач. Соискатель способен самостоятельно планировать исследование, осуществлять анализ полученных данных, формулировать выводы, творчески использовать в научной и деятельности знания фундаментальных и прикладных разделов молекулярной и классической генетики.

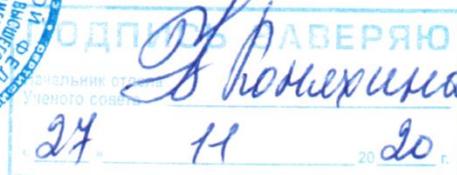
Диссертационная работа Романова Дмитрия Евгеньевича по своей научной новизне, теоретической и практической значимости соответствует требованиям п. 9 Положения «О порядке присуждения ученых степеней» от 24.09.2013 №842 (ред. от 01.10.2018, с изм. От 26.05.2020), предъявляемым к диссертациям, а ее автор заслуживает присуждения ему ученой степени кандидата биологических наук по специальности 03.02.07 – «Генетика».

Орлов Юрий Львович
Доктор биологических наук, профессор РАН,
Профессор,
Кафедра информационных и интернет-технологий
Институт цифровой медицины

Федеральное государственное автономное образовательное учреждение высшего образования Первый Московский государственный медицинский университет имени И.М. Сеченова Министерства здравоохранения Российской Федерации (Сеченовский Университет)

119991, Москва, ул. Трубецкая, д.8, стр. 2
Тел./факс: +7(499) 248-01-81
Эл.почта: y.orlov@sechenov.ru

27 ноября 2020 г.



Подпись Ю.Л. Орлова заверяю