

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 002.077.04
НА БАЗЕ ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО
УЧРЕЖДЕНИЯ НАУКИ ИНСТИТУТА ПРОБЛЕМ ПЕРЕДАЧИ
ИНФОРМАЦИИ ИМ. А.А. ХАРКЕВИЧА РОССИЙСКОЙ АКАДЕМИИ
НАУК ПО ДИССЕРТАЦИИ НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ
КАНДИДАТА БИОЛОГИЧЕСКИХ НАУК

аттестационное дело № _____

решение диссертационного совета
от 20 февраля 2017, протокол № 3

о присуждении Анне Анатольевне Гоглевой, гражданке Российской Федерации, ученой степени кандидата биологических наук.

Диссертация «Исследование CRISPR-систем прокариотического иммунитета методами сравнительной геномики» по специальности 03.01.09 – Математическая биология, биоинформатика, принята к защите 13 декабря 2016 года, протокол № 21 диссертационным советом Д 002.077.04 на базе Федерального государственного бюджетного учреждения науки Института проблем передачи информации им. А.А. Харкевича Российской академии наук, (127051, г. Москва, Большой Каретный переулок, д. 19, стр. 1., приказ о создании №978/нк от 16 декабря 2013 года).

Соискатель Анна Анатольевна Гоглева, гражданка Российской Федерации, 1987 года рождения, в 2008 году окончила с отличием биологический факультет Уральского Государственного Университета им. А.М. Горького, в по специальности «Биология». В 2010 году окончила с отличием магистратуру Пущинского Государственного Университета по специальности «Биология», направление «Микробиология и вирусология». В период подготовки диссертации, с 01.11.2010 по 01.11.2014 обучалась в заочной аспирантуре в отделе системной биологии и биоинформатики, группе биоинформатики Федерального государственного бюджетного учреждения науки Институте общей генетики им. Н.И. Вавилова Российской академии наук. С 14.05.2015 по настоящее время работает в Лаборатории Сейнсбери

Кэмбриджского Университета (Sainsbury Laboratory Cambridge University), в группе доктора Себастиана Шорнака на должности научного сотрудника.

Диссертация выполнена в отделе системной биологии и биоинформатики, группе биоинформатики Федерального государственного бюджетного учреждения науки Институте общей генетики им. Н.И. Вавилова Российской академии наук.

Научный руководитель - Ирена Игоревна Артамонова, кандидат биологических наук, старший научный сотрудник, руководитель группы биоинформатики в Федеральном государственном бюджетном учреждении науки Институт общей генетики им. Н.И. Вавилова Российской академии наук, старший научный сотрудник в Федеральном государственном бюджетном учреждении науки Институт проблем передачи информации им. А.А. Харкевича Российской академии наук.

Официальные оппоненты

1. Елена Николаевна Ильина, гражданка РФ, доктор биологических наук, профессор РАН, заместитель генерального директора по научной работе Федерального государственного бюджетного учреждения «Федеральный научно-клинический центр физико-химической медицины Федерального медико-биологического агентства».
2. Екатерина Евгеньевна Савицкая, гражданка РФ, кандидат биологических наук, старший научный сотрудник Автономной Некоммерческой Образовательной организации высшего образования «Сколковский Институт Науки и Технологий».

дали **положительные** отзывы на диссертацию.

Ведущая организация – Федеральное государственное автономное образовательное учреждение высшего образования "Балтийский федеральный университет имени Иммануила Канта" (БФУ им. И. Канта), г. Калининград в своем **положительном** заключении, подписанном кандидатом биологических наук Степаном Владимировичем Тоцаковым, зав. лабораторией геномики микроорганизмов ИЖС БФУ им. И. Канта и утвержденном доктором политических наук профессором Андреем Павловичем Клемешевым, ректором ФГАОУ ВО «БФУ им.И.Канта» указала ряд замечаний.

1. «... автор в своей работе не учитывает покрытия (coverage) используемых им контигов для того, чтобы установить взаимосвязь типа кассет и/или спейсеров с относительной представленностью (abundance) их хозяина. Анализ такой взаимосвязи был бы весьма интересен. Однако не исключено, что отсутствие такого анализа связано с тем, что информация о покрытии не присутствовала в используемых автором массивах данных».

2. «Недостаточно подробные подписи к рисункам и таблицам. Так, практически все рисунки раздела «Обзор литературы» не содержат дополнительных пояснений, хотя в целом в научной литературе принято давать подробные подписи к рисункам вне зависимости от их наглядности.»

3. «В рукописи присутствуют некоторые стилистические огрехи, связанные, как правило с не общепринятым переводом английских терминов. Так, к примеру, английский термин sequencing read, как правило, переводится на русский как «прочтение»».

Несмотря на замечания, в отзыве содержится высокая оценка научного уровня выполненной работы.

Соискатель имеет 6 статей в российских и международных рецензируемых журналах, включенных в перечень ВАК, из них 3 статьи по теме диссертации, общим объемом 27 страниц. Соискателем опубликовано 8 работ в материалах международных конференций.

Статьи по теме диссертации:

1. Gogleva AA, Gelfand MS, Artamonova II. Comparative analysis of CRISPR cassettes from the human gut metagenomic contigs // BMC Genomics. – 2014. – Mar 17.–15:202.

2. Гоглева А.А, Артамонова И.И. CRISPR системы: структура и гипотетические функции // Природа. – 2014. – № 1186. – С. 16 - 21.

3. Гоглева А.А, Артамонова И.И. CRISPR системы: механизм действия и применения // Природа. – 2014. – № 1186. – С. 3 — 9.

Тезисы докладов:

1. Gogleva A. Composition of CRISPR-cassettes reflects the virome diversity in the human gut microbiome / Gogleva A.A, Artamonova I.I. Труды 35-й конференции молодых ученых и специалистов ИППИ РАН (ИТиС'12), Петрозаводск, 2012. - С. 262 - 264.

2. Гоглева А. CRISPR-системы прокариотического иммунитета в микробиоме человека/ Гоглева А.А, Артамонова И.И. Труды 34-й конференции молодых ученых и специалистов ИППИ РАН (ИТиС'12), Геленджик, 2011. - С. 135 - 139.

3. Gogleva A. CRISPRome reflects the viriome composition in the human gut microbiome / Gogleva A.A, Artamonova I.I. CRISPR: Evolution, Mechanisms and Infection University of St Andrews, UK, 2013. - P017, P. 12

4. Gogleva A. CRISPRome reflects the viriome composition in the human gut microbiome / Proceedings of Russian-German seminar «Regulation and Evolution of Cellular Systems» (RECESS, Регуляция и эволюция клеточных процессов), Венеция, 2013.

5. Gogleva A. Artamonova I. CRISPR systems in microbiomes. Proceedings of Moscow Conference on Computational Molecular Biology (MCCMB'13), Москва, 2013.

Вклад диссертанта в опубликованные работы по теме диссертации состоит в непосредственном планировании исследований, участии в формулировках задач, теоретической разработке и практической реализации методов, обработке и анализе данных. Во всех трех публикациях диссертант является первым автором.

Полученные результаты могут быть использованы при чтении спецкурсов на биологических и биоинформатических факультетах ВУЗов, на курсах повышения квалификации биологов и/или медицинских работников. Также, полученные автором результаты и отработанные методические подходы могут служить базисом для проведения последующих исследований по динамике взаимодействий между вирусами и микробами, населяющими кишечник человека. Несомненно, отдельный интерес представляют дополнительные структурно-функциональные исследования новых CRISPR-ассоциированных белков, идентифицированных автором в составе CRISPR-кассет. Исследования такого рода проводятся в Институте биологии гена РАН, Институте молекулярной генетики РАН, Институте молекулярной биологии РАН, Институте общей генетики РАН, ФИЦ «Биоинженерия» РАН, а также в научных подразделениях многих федеральных университетов.

На диссертацию поступило 8 отзывов, из них — 5 отзывов на автореферат, все **отзывы положительные**. В отзывах указывается, что не вызывает сомнений актуальность темы рассматриваемой диссертации, а также научная и практическая ценность работы. Диссертация является законченным научно-исследовательским трудом, выполненным автором самостоятельно на высоком научном уровне. В ходе работы проанализирован состав CRISPR-кассет трёх метагеномных коллекций кишечника человека, двух из них — впервые. Подавляющая часть реконструированных кассет описана впервые. Определено происхождение спейсеров и подробно исследована динамика спейсеров и кассет внутри более чем 140 индивидуальных микробиомах человека, принадлежащих географически разобшенным популяциям. Показано что набор CRISPR-кассет и последовательность спейсеров внутри кассет очень специфичны, обнаружено лишь небольшое число общих спейсеров между индивидуальными метагеномами. Выявлено два функциональных класса спейсеров: спейсеры с мишенями и общие спейсеры. Показано что они статистически значимо сдвинуты в противоположных направлениях внутри кассеты: спейсеры с мишенями располагаются рядом с лидерной последовательностью, в то время как общие спейсеры сдвинуты к дистальному концу кассет и соответствуют более древнему состоянию CRISPR-иммунитета. Исследования являлись новыми на момент публикации, их достоверность подтверждена последующими работами. Результаты, полученные автором, являются новыми научными знаниями по специальности «03.01.09. - математическая биология, биоинформатика».

Таким образом, диссертационная работа Анны Анатольевны Гоглевой на тему «Исследование CRISPR-систем прокариотического иммунитета методами сравнительной геномики» является законченным научно-квалификационным исследованием и соответствует требованиям «Положения о порядке присуждения ученых степеней», утвержденного Постановлением Правительства РФ от 24 сентября 2013 г. №842.

В отзыве оппонента Елены Николаевны Ильиной, доктора биологических наук, профессора РАН Е, заместителя генерального директора

по научной работе Федерального государственного бюджетного учреждения «Федеральный научно-клинический центр физико-химической медицины Федерального медико-биологического агентства» отмечены достоинства диссертации, а также сделан ряд замечаний.

1. «Разделы 2.2.1 «Идентификация CRISPR-кассет, процедура фильтрации» и 2.2.4 «Определение происхождения спейсеров (поиск протоспейсеров)» содержат избыточную информацию, касающуюся уже полученных автором результатов, что отвлекает читателя от беспристрастного понимания применяемых алгоритмов поиска».

2. В отзыве отмечен ряд «неудачных формулировок положений - «исследование динамики спейсеров» и «контиги отнесены к типу *Firmicutes*», интерпретацию которых хочется услышать от автора».

В отзыве отмечено, что «существенных замечаний к работе нет, но есть один комментарий». «Как причину наблюдаемых отличий таксонов, доминирующих по содержанию CRISPR кассет, от таковых, рассчитанных на основании представленности генов 16S рПНК, автор приводит два источника счетных ошибок: различия в копиях гена 16S рПНК (число копий гена может варьировать от 1 до 15 в зависимости от вида) и разная эффективность амплификации генов 16S рПНК разных таксонов при использовании универсальных праймеров. Последний аргумент не применим в настоящем исследовании, поскольку в работе использовались данные метагеномных проектов, полученные *shotgun* секвенированием».

В отзыве оппонента Екатерины Евгеньевны Савицкой, кандидата биологических наук, старшего научного сотрудника Сколковского института науки и технологий в целом высоко оценившей диссертацию отмечен ряд недочетов.

1. «Литературный обзор дает достаточно подробное представление о состоянии исследований CRISPR-Cas систем, но, к сожалению, не освещает самые последние достижения. CRISPR область развивается чрезвычайно стремительно, в конце 2015 года К.Макаровой с коллегами (Makarova et al., 2015) была предложена новая классификация, которая сейчас и является общепринятой».

2. Присутствуют некоторая неточность формулировок (CRISPR-Cas системы иногда обозначаются как «CRISPR», иногда как «CRISPR-Cas», «CRISPR локус» в ряде случаев является синонимом CRISPR кассеты, а в ряде случаев обозначает *cas* гены, общепринятый термин «CRISPR-интерференция» не упоминается и заменен на «иммунный ответ»), неудачные словосочетания (например, «плазмидный геном»), ошибки пунктуации.

3. Классификация белков Cas приведена не полностью, также возможно стоило бы уделить большее внимание кратко затронутым в работе, но интересным и неоднозначным вопросам о селективном преимуществе CRISPR-Cas систем для клетки хозяина, роли горизонтального переноса генов в их распространении, вне-иммунным функциям CRISPR-Cas систем».

Также в отзыве отмечен ряд интересных результатов, полученных диссертантом «при поиске и анализе распределения протоспейсеров, соответствующих спейсерам выявленных CRISPR-кассет. 1) соответствие удалось выявить только менее чем для 1% спейсеров, что, возможно, свидетельствует об ограниченности наших представлений о вирусных геномах или, как альтернативный вариант, о других дополнительных способах возникновения новых спейсеров; 2) пары спейсер-протоспейсер «отталкиваются», то есть редко колокализуются в пределах микробиома одного индивидуума, подтверждая эффективность иммунной функции CRISPR-Cas систем; 3) спейсеры с мишенями статистически значимо сдвинуты к лидерному концу CRISPR-кассет, отсутствие мишеней более старых спейсеров, возможно, отражает очень высокую скорость эволюции бактериофагов».

В отзыве также отмечено, что полученные оригинальные данные и результаты «открывают возможности для дальнейшего изучения. «Например, было бы интересно узнать, насколько выявленные спейсеры пересекаются со спейсерами уже аннотированных (депонированных в базы данных) CRISPR-кассет, какова средняя длина повторов и спейсеров в выявленных CRISPR-кассетах, есть ли протоспейсеры среди последовательностей геномов бактерий, а не только в геномах бактериофагов и плазмид. Интересно было бы провести определение таксономической принадлежности контигов, не содержащих CRISPR-повторы, из этих же образцов и провести

сравнительный анализ с данными, полученными на основании контигов, содержащих CRISPR-кассеты, в результате которого, возможно, была бы получена информация о распространении CRISPR кассет в различных таксономических группах. Кроме того, разработанные в диссертационной работе Гоглевой А.А. алгоритмы могут быть применены в исследованиях CRISPR-Cas систем бактериальных сообществ других экологических ниш, что позволит приблизиться к пониманию их реального разнообразия.

Диссертационная работы Гоглевой А.А. проведена на большом объеме данных и содержит статистически достоверные результаты. Приведенные в настоящем отзыве замечания несколько не умаляют достоинства работы и не снижают высокую оценку в целом».

В отзыве на автореферат Якова Игоревича Давыдова, кандидата биологических наук сотрудника лаборатории эволюционной биоинформатики и лаборатории вычислительной филогенетики департамента экологии и эволюции Университета Лозанны отмечено, что «автореферат подробным образом излагает основные полученные результаты, а также кратко описывает использованные методы. Полученные результаты несомненно представляют научный интерес. Хочется отдельно отметить внимание соискателя к методологическим деталям, сочетающееся со способностью ясно видеть общую картину». В целом дана положительная оценка работе и перечислен ряд замечаний, которые «не снижают общей ценности диссертационной работы».

В отзыве на автореферат Дмитрия Сергеевича Карпова, кандидата биологических наук, старшего научного сотрудника Федерального государственного бюджетного учреждения науки Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, дана положительная оценка работе, существенных замечаний по работе нет.

В отзыве на автореферат Александра Вячеславовича Лаврова, кандидата медицинских наук, заведующего лабораторией мутагенеза Федерального государственного бюджетного учреждения науки Медико-генетического научного центра дана высокая оценка работе. В отзыве указано, что «Автореферат написан очень хорошим языком, понятно и грамотно. Отмечу только один недостаток: ДНК, контиги и т.п. измеряется

по-русски не в Мегабазах, а в миллионах оснований или нуклеотидов, т.е. мегануклеотидах. Мегабазы — сленг».

В отзыве на автореферат Михаила Алексеевича Пятницкого, кандидата биологических наук, старшего научного сотрудника Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича» отмечены следующие недостатки: «1) для анализа выбрано три метагеномных исследования, которые различаются между собой по объему выборки, возрасту, диете и применяемой технологии секвенирования. Соответственно сильно затрудняется перенос результатов и биологическая интерпретация сравнительного анализа. 2) Методы, использованные в работе, описаны недостаточно подробно. 3) Хотелось бы видеть больше сравнений с данными литературы». Отмечено, что указанные недостатки «не снижают высокой общей научно-практической ценности диссертационной работы».

В отзыве на автореферат Ильи Валерьевича Кубланова, кандидата биологических наук, старшего научного сотрудника Федерального исследовательского центра «Фундаментальные основы биотехнологий» Российской академии наук отмечено что автореферат «не содержит серьезных недостатков, но есть некоторые результаты, наблюдения и выводы, о которых можно было бы поговорить: 1) Из-за того, что были выбраны очень сложные по составу сообщества микроорганизмов, не удалось собрать большого количества контигов, содержащих кассету целиком. Возможно, имело смысл исследовать более простые сообщества; 2) Заниженное число CRISPR-содержащих контигов филума *Bacteroidetes*, на мой взгляд, довольно бездоказательно объясняется множественностью копий 16S рРНК у этих микроорганизмов, а также тем, что данные микроорганизмы могли бы полагаться на CRISPR-Cas системы других видов, что, по-моему, крайне маловероятно; 3) На мой взгляд, можно было бы провести более глубокий поиск протоспейсеров в геномах прокариот (в частности в их профагах и плазидах); 4) Вывод о высокой эффективности CRISPR-Cas систем все-таки выглядит несколько смелым с учетом всех технических особенностей обработки метагеномных ридов (несомненно, известных автору), ведущих, по-видимому, к большой потере вирусной компоненты». В заключении

указано, что «эти мелкие замечания и комментарии нисколько не снижают ценность данной диссертационной работы».

Выбор официальных оппонентов и ведущей организации обосновывается тем, что за последние годы ими было опубликовано большое количество научных работ, посвященных изучению CRISPR-Cas-систем и геномики микроорганизмов.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

проанализирован состав CRISPR-кассет трёх метагеномных коллекций кишечника человека, двух из них — впервые;

подавляющая часть реконструированных кассет **описана впервые**;

определено происхождение спейсеров и подробно исследована динамика спейсеров и кассет внутри более чем 140 индивидуальных микробиомов человека, принадлежащих географически разобшенным популяциям.

Определено таксономическое положение контигов, содержащих CRISPR-кассеты.

1. **Определен** тип CRISPR-Cas систем для идентифицированных кассет.

Теоретическая значимость исследования обоснована тем, что:

более 90% кассет идентифицированных в метагеномных данных микробиома человека описано **впервые**;

большая часть контигов, содержащих идентифицированные кассеты, была отнесена к типу *Firmicutes*, обнаружена только одна кассета в контиге архейного происхождения;

менее чем для 1% спейсеров протоспейсеры найдены среди известных последовательностей, задепонированных в коллекцию NR базы данных GenBank;

77% найденных **протоспейсеров** обнаружены в собственных метагеномных коллекциях, это наблюдение указывает на то, что истинное вирусное разнообразие до сих пор недоисследовано и недопредставлено в базах данных последовательностей;

показано, что набор CRISPR-кассет, как и последовательность спейсеров внутри кассет специфичны, обнаружено лишь небольшое число общих спейсеров между индивидуальными метагеномами, это наблюдение указывает на то, что даже при относительном сходстве видового состава микробиомов человека, цепь событий, приводящая к формированию CRISPR-опосредованного иммунитета уникальна для каждого конкретного индивидуального метагенома и активного обмена кассетами между микробиомами не просиходит;

показано, что спейсеры и прооспейсеры статистически значимо «отталкиваются», то есть не колокализуются в пределах индивидуальных микробиомах, возможно указывая на высокую эффективность CRISPR-Cas систем и быструю элиминацию вирусов из сообщества;

выявлены пары спейсер-протоспейсер, простирающиеся между индивидуальными микробиомами человека, в том числе из географически удаленных популяций, что может говорить в пользу существования общего пула широко распространённых (универсальных) вирусов, характерных для микробиома человека в целом;

выявлено два функциональных класса спейсеров: спейсеры с мишенями и общие спейсеры;

показано, что они статистически значимо сдвинуты в противоположных направлениях внутри кассеты: спейсеры с мишенями располагаются рядом с лидерной последовательностью, и отражают стадию адаптации CRISPR-Cas систем к новым вирусам, в то время как общие спейсеры сдвинуты к дистальному концу кассет и соответствуют более древнему состоянию CRISPR-иммунитета;

Значение полученных соискателем результатов исследования для практики подтверждается тем, что:

Было выявлено большое количество ранее не известных CRISPR-кассет и CRISPR-повторов, что свидетельствует о том, что реальное природное разнообразие CRISPR-Cas систем в значительной степени еще не исследовано. Полученные результаты важны для понимания фундаментальных аспектов эволюции прокариот и вирусов в рамках

симбиоза с человеком, а также для возможного развития новых методов и подходов к диагностике и прогнозу заболеваний человека.

Свидетельства достоверности результатов исследования:

- таксономическое распределение CRISPR-содержащих контигов для двух метагеномов (HMP и DG) качественно совпадает с результатами анализа генов 16S рРНК.

- статистическая достоверность гипотезы о независимом распределении спейсеров и протоспейсеров в пределах индивидуальных метагеномов проверена при помощи статистического теста, аналогичного тесту Кохрана-Мантеля-Ханцеля для наблюдаемых данных и 100'000 симулированных наборов данных.

- для оценки того, насколько значимо наблюдаемое число общих спейсеров между индивидуальными метегеномами применили моделирование на основе случайных перестановок.

- для оценки значимости смещения спейсеров с мишенями к лидерному концу CRISPR-кассет, а общих спейсеров — к дистальному концу, применяли моделирование по методу Монте-Карло.

Личный вклад соискателя состоит в планировании исследований, участии в постановке задач, теоретической разработке и практической реализации методов, анализе данных и написании статей. Результаты, описанные в диссертации, получены автором самостоятельно. В диссертации решены задачи, имеющие значение для развития исследований в области исследования эволюции CRISPR-Cas систем прокариот, работа обладает внутренним единством, что подтверждается наличием последовательного плана исследования и непротиворечивыми выводами. По своему содержанию диссертация отвечает паспорту специальности 03.01.09 – Математическая биология, биоинформатика.

Диссертационный совет пришёл к выводу о том, что диссертация представляет собой завершённую научно-квалификационную работу, имеющую теоретическое и практическое значение для медицины и биотехнологии. По актуальности, новизне, практической значимости диссертация соответствует требованиям, установленным «Положением о порядке присуждения ученых степеней», утвержденным постановлением

Правительства Российской Федерации от 24 сентября 2014 года No 842, предъявляемым к диссертациям на соискание ученой степени кандидата биологических наук.

На заседании 20 февраля 2016 года диссертационный совет принял решение присудить Анне Анатольевне Гоглевой ученую степень кандидата биологических наук по специальности 03.01.09 – Математическая биология, биоинформатика.

При проведении тайного голосования диссертационный совет в количестве 15 человек, из них 8 докторов наук по специальности рассматриваемой диссертации, участвовавших в заседании, из 21 человека, входящих в состав совета, проголосовали за - 14, против - 0, недействительных бюллетеней - 1.

Председатель

диссертационного совета Д 002.077.04

д.б.н., профессор


М.С. Гельфанд

Ученый секретарь

диссертационного совета Д 002.077.04

д.б.н., профессор

20 февраля 2017 г.


Г.И. Рожкова