



Сколковский институт науки и технологий

Автономная некоммерческая образовательная
организация высшего образования
«Сколковский институт науки и технологий»
143025, Московская область, Одинцовский район,
дер. Сколково, ул. Новая, дом 100
ОГРН 1115000005922
ИНН/КПП 5032998454/503201001
Тел.: +7 (495) 280-14-81

Отзыв официального оппонента

на диссертационную работу Гоглевой Анны Анатольевны «Исследование CRISPR-систем прокариотического иммунитета методами сравнительной геномики», представленную на соискание ученой степени кандидата биологических наук по специальности 03.01.09 – «математическая биология и биоинформатика»

Диссертационная работа А.А. Гоглевой посвящена изучению разнообразия и динамики CRISPR-Cas -систем адаптивного иммунитета прокариот в микробиоме человека. Принцип действия CRISPR-Cas систем основан на встраивании фрагмента генома чужеродного агента в локус, называемый CRISPR-кассетой; последующая экспрессия приобретенного фрагмента в составе короткой РНК обеспечивает узнавание и внесение разрыва в комплементарную, т.е. чужеродную ДНК. Молекулярные механизмы действия CRISPR-Cas систем, обеспечивающие специфическое узнавание и расщепление ДНК мишени, легли в основу создания новых высокоэффективных инструментов редактирования геномов про- и эукариот и обеспечили лавинообразный рост публикаций, посвященных CRISPR-Cas системам. Исследования ведутся на стыке многих областей науки: молекулярной биологии, генетики, микробиологии, однако, без преувеличения можно сказать, что решающую роль играют методы биоинформатики, в первую очередь сравнительной геномики. В 2005 году при помощи биоинформатических методов было замечено, что некоторые спейсеры CRISPR-кассет соответствуют участкам геномов бактериофагов и плазмид, названных в последствие протоспейсерами, что привело к предположению об иммунной функции CRISPR-Cas систем. Затем именно при помощи биоинформатического поиска были предсказаны многие новые типы и классы CRISPR-Cas систем, и стала возможна их

классификация. Биоинформатический анализ набора спейсеров CRISPR кассет, их соответствия протоспейсерами-мишеням потенциально способен выявить большой пласт информации о взаимоотношениях вирусов и бактерий в пределах микробиома, динамике микробных сообществ во времени, географическом распространении и т.д. В результате CRISPR-Cas системы могут стать мощными инструментами для эпидемиологических, экологических и популяционных исследований. Примеры работ, посвященных разнообразию спейсеров и их соответствию протоспейсерам в различных экологических нишах, начинают появляться в литературе, судя по их результатам, исследования в данном направлении будут стремительно развиваться. Работу Гоглевой А.А., посвященную изучению разнообразия и динамики CRISPR-Cas систем, несомненно, находится в рамках данного актуального и перспективного направления.

В диссертационной работе А.А.Гоглевой ставятся задачи изучить разнообразие и таксономическую принадлежность CRISPR-Cas систем на основе метагеномных данных, определить состав спейсеров и произвести поиск соответствующих им протоспейсеров, провести сравнительный анализ различных наборов данных одной экологической ниши. Объект исследования удачно выбран: микробиом кишечника человека, представляющий особый интерес с прикладной медицинской точки зрения. В ходе выполнения исследования был проанализирован большой массив данных, составляющий более 2 млн контигов, доступных на момент выполнения анализа, получены интересные и новые результаты, характеризующие микробиом кишечника человека с точки зрения функционирования CRISPR-Cas систем.

Диссертационная работа изложена на 101 странице, содержит 17 рисунков и 7 таблиц, список литературы содержит 162 источника. Текст работы разбит на 4 главы: «Обзор литературы», «Данные и алгоритмы», «Результаты и обсуждение», «Гипотезы и перспективы». Текст автореферата соответствует содержанию диссертации.

Обзор литературы состоит из двух частей. Первая часть посвящена CRISPR-Cas-системам, вторая – проблемам микробиома человека, а также роли CRISPR-Cas систем в исследованиях микробиома человека. 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. К сожалению, автор приводит и придерживается в своей работе старой

классификации CRISPR-Cas систем. Следует также заметить, что присутствует некоторая неточность формулировок (CRISPR-Cas системы иногда обозначаются как «CRISPR», иногда как «CRISPR-Cas», «CRISPR locus» 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»), встречаются неудачные словосочетания (например, «плазмидный геном») и ошибки пунктуации. 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.

В главе «Данные и алгоритмы» подробно охарактеризованы наборы данных, использованных в работе, а также описаны примененные алгоритмы, как заимствованные, так и разработанные автором, аргументирован их выбор. Глава сопровождается ссылками на источники данных, таблицами и схемой примененных алгоритмов, значительно облегчающими восприятие материала.

Описание результатов диссертационной работы приводится в главе 3. В наборах метагеномных данных микробиома человека с помощью разработанного алгоритма были выделены CRISPR-кассеты. Несомненным преимуществом работы является сочетание нескольких инструментов обнаружения CRISPR-кассет (CRISPR finder, PILER-CR, CRT), позволяющих максимально избежать ошибок, свойственных каждому отдельно взятому подходу. В результате поиска только в данном типе микробиома, а именно микробиоме кишечника человека, 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. Следующими этапами работы стали определение принадлежности CRISPR кассет к таксономической группе организмов исходя из анализа фланкирующих генов, не относящихся к CRISPR-Cas системам, и отнесение к определенному типу CRISPR-Cas систем путем выявления гомологии прилежащих *cas* генов с *cas* генами уже классифицированных CRISPR-Cas систем. Для значительной доли CRISPR-кассет удалось провести выявление таксономической группы на уровне домена, в некоторых случаях удалось установить принадлежность CRISPR-кассеты к определенному типу CRISPR-Cas систем. При поиске и анализе распределения протоспейсеров,

соответствующих спейсерам выявленных CRISPR-кассет, были получены следующие интересные результаты: 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, имеющие соответствующие протоспейсеры, статистически значимо сдвинуты к лидерному концу CRISPR-кассет, отсутствие мишеней более старых спейсеров, возможно, отражает очень высокую скорость эволюции бактериофагов.

В целом, в ходе данного исследования были получены новые оригинальные данные и результаты, которые открывают направления 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 в 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. Выводы представляются обоснованными и актуальными. Результаты работы опубликованы в иностранных и российских реферируемых журналах, рекомендованных ВАК.

Диссертационная работы А.А. Гоглевой «Исследование CRISPR-систем прокариотического иммунитета методами сравнительной геномики» является самостоятельным законченным исследованием, выполненным на высоком профессиональном уровне, по научной и практической

значимости, методическому уровню полностью удовлетворяет требованиям пункта 9 «Положения о порядке присуждения ученых степеней», утвержденных постановлением Правительства Российской Федерации №842 от 24 сентября 2013 г. (с изменениями в редакции постановлений Правительства Российской Федерации №335 от 21.04.2016 г., №478 от 02.08.2016 г.), предъявляемым к кандидатским диссертациям. Гоглева А.А., несомненно, заслуживает присуждения ученой степени кандидата биологических наук по специальности 03.01.09 «математическая биология, биоинформатика».

Савицкая Екатерина Евгеньевна,
кандидат биологических наук,
старший научный сотрудник,
Автономная некоммерческая образовательная организация
высшего образования
«Сколковский институт науки и технологий»,

savitskaya.e@yandex.ru

тел. +7(916)4939857

02» февраля 2017 г.

Савицкая

