

ОТЗЫВ

официального оппонента на диссертационную работу Цой Ольги Владиславовны на тему «Эволюция систем регуляции транскрипции в геномах бактерий», представленную на соискание ученой степени кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика.

Работа Ольги Владиславовны представляет собой биоинформатическое исследование эволюции систем регуляции транскрипции у бактерий. Полноценные эволюционные исследования, проводимые в последние годы, стали возможны именно благодаря синергии методов высокопроизводительного прочтения ДНК и методов биоинформатики, позволяющим анализировать объемные массивы данных. В этом смысле тематика представленной работы полностью соответствует современному развитию исследовательской области, фокусируясь на конкретных бактериальных регуляторных системах на уровне транскрипции. С одной стороны, именно сегодня объем накопленных для различных бактерий геномных данных позволяет проводить полноценный статистический анализ и в полную силу использовать методы сравнительной геномики. С другой стороны, регуляторные системы, включающие факторы транскрипции, у прокариот устроены более прямолинейно по сравнению с эукариотами, и это дает возможность более явно проследить содержательные эффекты, в том числе, уверенно выйти на уровень устойчивых элементов регуляторных сетей, что и сделано автором рассматриваемой работы. Первая часть работы, посвященная регуляторному пути деградации этаноламина, в некотором смысле следует классической схеме, а именно описана структура соответствующего оперона и проанализировано филогенетическое дерево различных бактерий, способных утилизировать этаноламин. Особенно замечательно, на мой взгляд, что авторам с помощью методов анализа последовательностей удалось выявить не известный ранее мотив ДНК-белкового узнавания транскрипционного фактора EutR и провести непрямую валидацию этого мотива путем анализа потенциальных генов-мишеней. Вторая часть работы, на мой взгляд, поднимает исследование на принципиально новый уровень: от исследования достаточно

конкретной и ограниченной регуляторной сети авторы переходят к анализу фундаментальных особенностей устройства и эволюции регуляторных сетей, так называемых мотивов в генных сетях (сетевых мотивов). В рамках этой парадигмы проведен детальный анализ эволюции базовых сетевых мотивов, как для близких штаммов *Escherichia coli*, так и на уровне порядка *Enterobacteriales*.

У меня нет никаких серьезных замечаний к научной стороне работы. Более того, часто страдающее у кандидатских диссертаций изложение материала в данном случае соответствует очень высокому уровню. Тем не менее, тексту работы и общему оформлению недостает лоска и завершенности структуры и формулировок. Приведу ряд конкретных примеров. В «целях и задачах исследования» автор пишет об «эволюции регуляторных блоков на примере блока треугольник». На мой взгляд сам термин «блок треугольник» не является общепринятым и нигде до этого не обсуждается; таким образом в этот момент читателю трудно понять, о чем на самом деле пойдет речь. В принципе, можно было предложить и более академический вариант перевода классического сетевого мотива «feed-forward loop», например, «петля прямой связи». В разделе, посвященном научной новизне, автор называет мотив участка связывания регуляторным элементом в метаболическом пути. На мой взгляд, элементами пути являются гены и соответствующие белки, а мотив участка связывания – лишь способ представления ДНК-белкового узнавания. Присутствуют и неудачные формулировки, такие как «экспоненциально растущее количество геномов» (имеется в виду количество прочитанных геномов), и «разные методы сравнительной геномики». В положениях, выносимых на защиту, фигурируют взаимодействия типа «Л – стрелочка – ген», было бы более грамотно явно описать тип взаимодействий. В теме работы явно звучит тот факт, что работа посвящена бактериям, но далее во введении при описании регуляции транскрипции было бы правильно еще раз озвучить, что речь идет именно про прокариот. Еще одна замечательная фраза, видимо напрямую пришедшая из технического научного жаргона, «если ген имеет участок связывания транскрипционного фактора в одном геноме...» (стр. 10 диссертации). Там же начинается описание анализа сайтов связывания, и оно лишено строгости и научной аккуратности. Например, делается спорное утверждение, «одним из

наиболее эффективных является матрица позиционных весов». Действительно, сегодня это наиболее массово используемый метод, но его «наибольшая эффективность» по современным данным достаточно сомнительна. Там же речь идет о «консервативности самой позиции», и, поскольку модели сайтов связывания описаны недостаточно детально, из контекста может быть не до конца понятно о какой консервативности (информационной или эволюционной) идет речь. Также оставлен в стороне важный технический вопрос выбора порогов для предсказаний сайтов связывания, например «сильное уменьшение .. веса» (стр. 67) явно не формализовано. Описание метода филогенетического футпринтинга также лишено строгости, например, в нем присутствует загадочная фраза «до тех пор, пока выравнивание имеет смысл». На странице 16 упоминаются «гены двухкомпонентной регуляторной системы», при этом не очень понятно, какая именно специфическая система имеется в виду в данном случае. На странице 29 автор говорит о критерии «лучшего двухстороннего поиска» (bidirectional best hit), вероятно перевод также не является общепринятым. В обсуждении автор выдвигает, но недостаточно подкрепляет смелый и общий тезис: «именно мутации являются основным инструментом эволюции регуляторных взаимодействий». Следовало бы подкрепить эту мысль, если она действительно кажется автору важной, более детальным описанием и ссылками на литературные источники.

Заметной части раздела 3.2, посвященной анализу сайтов связывания, не хватает ясности: следовало расположить выравнивания и соответствующие визуализации мотивов на одном рисунке или близко в тексте в одном порядке, явно указать число последовательностей, использованных в каждом случае для построения мотива. Аналогичная проблема плохого структурирования материала возникает в разделе 4.4., где по таблицам 11-14 трудно понять, какие именно значения из каких именно колонок сравнивались при подсчете статистической значимости. Явное приведение таблиц сопряженности заметно упростило бы восприятие результатов. Кроме того, с учетом малого размера выборок, было бы более уместно использовать точный тест Фишера, а не критерий хи-квадрат.

Ряд претензий возникает и в связи с иллюстрациями. Рисунок 1 иллюстрирует метаболосому в клетке, но достаточно трудно понять, как именно она выделена на достаточно

мутной микрофотографии. Также не до конца понятно, какую информацию, важную для работы, этот рисунок несет, какую именно мысль иллюстрирует. Рисунки 19-22 лишены подписей осей.

Наконец, встречаются и опечатки, например «поиск новых *регулируемых* им генов» (стр. 17) и «исчезновение регуляторного гена» (стр. 67, имелся в виду регулируемый ген).

Указанные недостатки, хотя и усложняют чтение и понимание работы, никоим образом не снижают ее научного уровня. Выводы и научные положения, сделанные в работе, полноценно обоснованы, предложен ряд важных и новых идей; личный вклад соискателя в работу несомненен, в том числе, во всех публикациях по теме работы Ольга Владиславовна является первым автором. Автореферат полноценно отражает содержание диссертации. Таким образом, в заключение считаю необходимым отметить, что работа Ольги Владиславовны Цой полностью соответствует требованиям «Положения о присуждении ученых степеней», утвержденным Постановлением № 842 Правительства РФ от 24 сентября 2013 г. В свою очередь, автор безусловно заслуживает присуждения ученой степени кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика.

кандидат физико-математических наук,
старший научный сотрудник
Лаборатории вычислительных
методов системной биологии
Федерального государственного
бюджетного учреждения науки
Института молекулярной биологии
имени В.А. Энгельгардта
Российской академии наук

 / И.В. Кулаковский /

14 октября 2014 года

