

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

**на диссертационную работу Цепилова Якова Александровича на тему
«Разработка и применение новых моделей в полногеномном анализе
ассоциаций», представленная на соискание ученой степени кандидата
биологических наук по специальности 03.02.07 –генетика**

Представленная к рассмотрению работа посвящена новым методам обработки данных полногеномного анализа ассоциаций (GWAS) и реализации предложенных методов на крупных GWAS, связывающих геномный полиморфизм и уровень метаболитов в плазме крови человека. Введение геномных и постгеномных технологий резко усилило значимость обработки данных, ведь без специальных программных инструментов их интерпретация человеческим сознанием невозможна. И наоборот, некорректное применение статистических методов в так называемых «больших» данных способно породить различные ошибки, в том числе ложноположительные результаты, особенно опасные в генетике человека. Это определяет общую актуальность заданного автором направления.

Основная цель работы – создать инструмент, определяющий модель наследования у каждого отдельного участка полиморфизма, преодолевшего порог достоверной ассоциации с заболеванием. Аддитивный или неаддитивный характер наследования кажется чем-то несложным: вспоминаются картинки из школьного учебника старших классов, где изображены сморщенные или гладкие семена гороха или постепенно меняющие цвет венчика растения ночной красавицы. Но ситуация требует серьезной статистической работы, когда имеется более миллиона переменных (участков геномного полиморфизма), несколько тысяч обследуемых и несколько сотен количественных признаков, с которыми требуется выявить достоверные ассоциации.

По умолчанию, при анализе ассоциации участков геномного полиморфизма с фенотипическими признаками, вклад в которые лишь частично связан с генотипом (например, распространенные заболевания человека), наследование считается аддитивным. Это, в общем, вытекает из общих соображений. Например, если участок полиморфизма приводит к миссенс-мутации, в результате чего мутантный фермент существенно меняет кинетические и/или термодинамические параметры

катализируемой реакции, мы вправе ожидать неаддитивного наследования. Этот механизм доминирования понимали еще в 80-е годы прошлого века (например, Kacser and Burns, 1981). Но такие случаи тщательно отслеживаются эволюцией, ведь изменения будут столь существенны, что будет наблюдаться патология с менделевским типом наследования, или такой индивид просто не выживет. Так что большинство участков однонуклеотидного полиморфизма (SNP), исследуемого в GWAS, оказывают слабое воздействие на признаки с аддитивным характером наследования. Однако, всегда есть потребность ранжировать такие участки. Чем больше указаний на неаддитивный характер наследования, тем ближе мы к раскрытию механизма воздействия, тем важнее выявленная ассоциация к диагностике, прогнозу или управлению заболеванием. Вот для чего, по сути, затевалось это исследование, эту сложную задачу поставил себе автор – разработать инструмент для выявления таких неаддитивных SNP в океанах информации современных GWAS. И надо сказать, что он с этим отлично справился.

Для неаддитивных SNP потребовалось создать специальную процедуру геномного контроля результатов GWAS. Автор изощренно и скрупулезно использовал статистическое моделирование для определения параметров геномного контроля в случаях кодоминирования, рецессивного и доминантного характера наследования, сверхдоминирования. Подход был проверен на данных GWAS, предоставляющих ассоциации с уровнем 150 метаболитов плазмы крови человека. В итоге было выявлено несколько SNP, которым больше подходит неаддитивная модель наследования. Для некоторых из них удавалось, хотя бы приблизительно, интерпретировать биохимические механизмы воздействия, что уже является существенным успехом.

Работа построена по традиционному плану. Обзор литературы выдает вовлеченность автора в развитие данной области знаний в мире. Он читается с интересом и может быть рекомендован для ознакомления с эволюцией мысли по изучению характера наследования, что в итоге привело к обработке больших данных GWAS. Все методы и результаты также получены в контексте актуальных исследований, ведущихся в мире, в рамках интенсивного международного

сотрудничества. Выводы адекватны полученным данным. Общее содержание работы полностью отражено в автореферате. Результаты работы позволяют приоритизировать результаты GWAS, что в перспективе имеет практическое значения для разработки диагностических и прогностических панелей.

Вместе с тем, к работе имеются некоторые замечания, не умаляющие, впрочем, сложившегося от знакомства с ней благоприятного впечатления.

Во-первых, мало внимания уделено методу оценки количественных признаков, с которыми выявляются ассоциации. Метаболиты измеряли посредством хромато-масс-спектрометра типа тройного квадрупольа с использованием готового набора для измерения 150 метаболитов AbsoluteIDQ p150. Известно, что аналитические характеристики, такие как погрешность измерения, при мультимплексном анализе падают. Однако, в работе не указано, каковы были значения погрешности измерений, коэффициента вариации оценки признаков. Я понимаю, что автор занимался обработкой данных и далек от процедуры их получения. Тем не менее, хотелось бы знать, каким образом при расчетах может быть учтена аналитическая точность оценки количественных признаков.

Кроме того, в работе отмечается некоторое невнимание к ее биохимическому смыслу. Да, SNP с неаддитивным типом наследования прокомментированы с точки зрения связи с метаболитами плазмы крови. Однако, ни в одной из таблиц в диссертации, не говоря об автореферате, не были детально раскрыты сами метаболиты, их химическая формула или полное название. Они обозначены кодами, расшифровать которые часто нельзя, не прибегнув к первоисточнику. Погрузившись в сложные расчеты, автор как будто на время забыл, что занимается биологией.

В тексте не обошлось без опечаток и стилистических ошибок. Обращает на себя внимание, что автор по всему тексту называет субъектов исследования «особями». Представляется, что в традициях русского языка в медицинской генетике принято все же называть людей индивидами или субъектами.

В заключение отмечу, что данная работа по степени вовлеченности в проблему, научной актуальности, значимости результатов существенно

превосходит средний уровень кандидатской диссертации. Требования по публикациям по теме работы успешно выполнены. Личный вклад автора не вызывает сомнений.

Заключение

Диссертационная работа Цепилова Якова Александровича на тему «Разработка и применение новых моделей в полногеномном анализе ассоциаций», представленная на соискание ученой степени кандидата биологических наук по специальности 03.02.07 – генетика, является законченным научно-квалификационным исследованием. По степени научной новизны, актуальности и практической значимости результатов диссертационна работа Я.А. Цепилова соответствует требованиям, предъявляемым к диссертациям на соискание степени кандидата биологических наук в соответствии с п.9 «Положения о порядке присуждения ученых степеней», утвержденного Постановлением Правительства РФ №842 от 24.09.2013 года, а автор заслуживает искомой ученой степени кандидата биологических наук по специальности 03.02.07 – генетика.

Зав. отделом протеомных исследований
Федерального государственного бюджетного научного учреждения
«Научно-исследовательский институт
биомедицинской химии имени В.Н. Ореховича» (ИБМХ),
зав. каф. биохимии Медико-биологического факультета
ГБОУ ВПО Российский национальный исследовательский
медицинский университет им. Н.И. Пирогова Минздрава России

Д-р биол. наук, проф. РАН

Мошковский Сергей Александрович

Адрес: ИБМХ, 119121, Россия Москва, ул. Погодинская, д. 10, стр.8, тел. +7 (499) 246-69-80, +7 (499) 246-34-66, факс: +7 (499) 245-08-57, эл. почта inst@ibmc.msk.ru, smosh@mail.ru

Подпись Мошковского
заверяю
Ученый секретарь ИБМХ к.х.н. Карпова Е.А.

