

ОТЗЫВ

официального оппонента на диссертацию А.А. Аракеляна "РАЗРАБОТКА БИОИНФОРМАТИЧЕСКИХ ПОДХОДОВ ДЛЯ ИЗУЧЕНИЯ МОЛЕКУЛЯРНЫХ МЕХАНИЗМОВ РАЗВИТИЯ ХРОНИЧЕСКИХ ЗАБОЛЕВАНИЙ ЧЕЛОВЕКА", представленной на соискание ученой степени доктора биологических наук по специальности 03.00.03 - "Молекулярная и клеточная биология".

Хорошо известно, что каждый год смертность от неинфекционных заболеваний (НИЗ), составляет большую часть от всех случаев смертей в мире. Поэтому борьба с НИЗ являются актуальным и приоритетным направлением биомедицины и смежных научных дисциплин. Отметим, что НИЗ развиваются в результате воздействия комбинации генетических, физиологических, экологических и поведенческих факторов, что неизбежно приводит к возникновению целого ряда проблем при прогнозировании и диагностике НИЗ, а также при их лечении. Многочисленные медико-биологические исследования показали, что возникшие проблемы удастся более или менее эффективно решать только тогда, когда выявлены молекулярные механизмы, лежащие в их основе. В настоящее время пожалуй только биоинформатический подход позволяет в достаточно полном объеме исследовать молекулярные механизмы развития хронических заболеваний. Именно этой важной и актуальной теме и посвящена докторская диссертация А.А. Аракеляна. Следует отметить, что любой нормально функционирующий организм, а также развитие патологий является результатом чрезвычайно сложноорганизованной сети физико-химических процессов. Так что имеющиеся биоинформатические идеи и подходы, хорошо зарекомендовавшие себя при анализе определенных заболеваний, оказываются мало эффективными, а иногда и непригодными при их использовании для анализа других заболеваний. А.А. Аракелян в диссертационной работе модифицировал и дополнил известные, а также разработал целый ряд новых более эффективных и универсальных биоинформатических подходов для изучения молекулярных механизмов развития ряда хронических заболеваний человека. Результаты, полученные в работе А.А. Аракеляна, открывают новые возможности исследования молекулярных механизмов заболеваний.

Диссертация состоит из введения, списков сокращений и основных обозначений, четырех глав, заключения и выводов. Общий объем диссертации 231 страниц, включая 63 рисунка и 16 таблиц. Список литературы содержит 428 источников.

Диссертация начинается с списка использованных сокращений и списка символов и обозначений.

Во введении обоснована актуальность проблемы, очерчен круг вопросов, рассмотренных в диссертации, сформулирована цель диссертационной работы. Приводится научная новизна диссертационной работы, ее теоретическая и практическая ценность, дана общая характеристика работы.

Глава 1 посвящена обзору литературы по теме диссертации. Приводится краткий обзор работ о современном состоянии исследования по молекулярным механизмам развития неинфекционных заболеваний человека. Обзор показывает, что одним из основных инструментов изучения молекулярных механизмов развития заболеваний, поиска лекарственных препаратов и диагностических маркеров является глобальная экспрессия генов. Приведены современные данные указывающие на то, что исследования на уровне транскриптома важны для интерпретации функциональных элементов генома и выявления молекулярных механизмов, лежащих в основе как нормальной деятельности клеток, так и развития патологических процессов. Обсуждены работы, где содержатся сведения о высокопроизводительных методах измерения генома, позволяющих получать большой объем информации, анализ которого неизбежно диктует разработку новых, более эффективных биоинформатических подходов и методов.

Глава 2 посвящена разработке новых биоинформатических подходов для анализа экспрессии генов. Ключевой задачей является определение тех генов и биологических процессов, которые ответственны либо за фенотипические различия между классами, либо же являются его последствиями. Частично эту задачу можно решить методами математической статистики. В диссертации предложен вариант решения этой задачи в общем виде, свободный от ряда допущений, лежащих в основе применения математической статистики. В основе подхода диссертанта лежит идея о логико-комбинаторном машинном обучении, разработанный в области распознавания образов. На пути реализации этой идеи, диссертанту удалось решить целый ряд проблем с учетом их биологической специфики, которые подробно описаны в диссертационной работе. Отмечу лишь некоторые из них: уменьшение вычислительной сложности алгоритмов за счет определения системы опорных множеств и создание алгоритма растущих опорных множеств, создание классификатора с использованием обучающей выборки, оценка ошибки/точности классификации, идентификация экспрессируемых генов на фоне «случайных» флуктуаций остального транскриптома и т.д.

В итоге был разработан комплекс алгоритмов и программных пакетов для анализа и интерпретации данных о глобальной экспрессии генов, который был успешно использован при изучении молекулярных механизмов широкого спектра хронических заболеваний человека.

Глава 3 посвящена разработке биоинформатических методов для моделирования активности биологических путей. Биологические пути представляют собой одну из фундаментальных основ жизнедеятельности клетки и организма, они образуют сеть физико-химических процессов, которые обеспечивают нормальную жизнедеятельность клеток и организма. Исследования биологических путей являются одной из наиболее актуальных задач современной биоинформатики и они направлены на выявление молекулярных механизмов развития заболеваний, определение мишеней для разработки лекарств. Следует заметить, что с появлением методов транскриптомики, протеомики и метаболомики концепция биологических путей позволила перейти от анализа на геномном уровне на уровень системной биологии. Появились принципиально новые возможности для понимания биологических процессов, протекающих в клетке в норме и патологии. Были разработаны ряд математических алгоритмов и компьютерных программ для анализа биологических путей. Среди множества алгоритмов и программных пакетов одним из первых для определения активности биологических путей, основанных на топологической информации и уровнях экспрессии генов и (или) белков был алгоритм Pathway Impact Analysis. В настоящее время одна из наиболее полно аннотированных баз данных биологических путей является KEGG Pathway. Несмотря на широкий спектр алгоритмов и программных пакетов для анализа биологических путей их практическая реализация сталкивается с различного рода ограничениями сильно уменьшая их рабочую эффективность. По этой причине диссертантом предложены программные пакеты для редактирования и визуализации биологических путей KEGG Pathways, в которые впервые внедрена возможность полуавтоматической коррекции взаимодействий между компонентами биологических путей, модификации по тканевой специфичности и белок-белковым взаимодействиям. Разработанные диссертантом программы KEGGParser и CyKEGGParser позволяют создавать библиотеки биологических путей на основе информации, хранящейся в базе данных KEGG Pathway. Разработан также алгоритм PSF, который позволяет оценить изменения активности биологических путей, основываясь на концентрации биомолекул - компонентов пути и взаимодействиях между ними. Предложенный в работе пакет программ существенно повысил точность результатов биоинформатического анализа и их интерпретацию. Следует отметить, что в диссертационной работе предложены также весьма важные разработки

алгоритмов и программных пакетов, позволяющих осуществить непосредственный переход от по-генного метода анализа дифференциальной экспрессии генов к анализу на уровне сигнальных путей. Эти разработки значительно улучшают точность определения биологических процессов, вовлеченных в развитие патологических состояний.

Глава 4 посвящена исследованию молекулярных механизмов патогенеза комплексных заболеваний человека. В этой главе представлены результаты исследований молекулярных механизмов моно- и полигенных заболеваний с помощью разработанных диссертантом биоинформатических алгоритмов и программных пакетов, при этом, при необходимости, были использованы интернет доступные данные по глобальной экспрессии генов. Был исследован целый ряд неинфекционных заболеваний человека и получены результаты, которые, несомненно, представляют интерес для практической биомедицины. Диссертант убедительно показывает, что биоинформатический подход, в основе которого лежит анализ молекулярных механизмов развития патологий может быть весьма эффективным. Полученные результаты при исследовании конкретных заболеваний продемонстрировали преимущества применения системных подходов для функционального анализа глобальной экспрессии генов, учитывающей особенности топологии и характер взаимодействий между генами в биологических путях, по сравнению с по-генным анализом дифференциальной экспрессии генов. Выявлена зависимость резистентности биологических путей к мутациям и идентифицированы ключевые гены, мутации в которых могут существенно повлиять на нормальные функции клеток. В ряде случаев, проведенные диссертантом исследования позволили идентифицировать ключевые гены и биологические пути, задействованные в развитии исследуемых заболеваний. Кроме того, было показано также, что системный анализ на уровне пертурбаций биологических путей является гораздо более информативным подходом исследования патогенеза заболеваний, чем по-генный анализ экспрессии генов. В этой главе показано, что сигнальные и метаболические пути устойчивы к мутациям, влияющим на белок-белковые взаимодействия благодаря множественным дублирующим ответвлениям и разветвленным топологиям. Показано также, что метаболические пути могут содержать узлы-концентраторы и «узкие» места, мутации в которых могут существенно влиять на общую активность данного пути.

Полученные результаты существенно обогащают современные представления о патогенезе хронических и онкологических заболеваний легких, посттравматического стрессового расстройства, аутоиммунных и аутовоспалительных заболеваний, открывают новые пути их диагностики и лечения.

Диссертация хорошо оформлена и иллюстрирована рисунками. Мои замечания к диссертации следующие.

1. Проблема выделения полезного сигнала на фоне случайных флуктуаций решается в каждом конкретном случае исходя из отношения сигнал/шум и минимальных априорных предположениях относительно формы полезного сигнала и характеристик шума. Из каких соображений был выбран метод выделения полезного сигнала, использованный в диссертации.

2. В подавляюще большинстве случаев сложные динамические процессы в биологических системах в которых имеются «узкие места» успешно моделируются аналитически с помощью дифференциальных уравнений. Из каких соображений в диссертации выбран путь численного моделирования.

К сожалению в диссертации и автореферате имеются некоторые опечатки и неточности, которые, однако, не искажают общего весьма положительного впечатления от работы. В целом, диссертационная работа А.А. Аракеляна выполнена на высоком научном уровне и вносит значительный вклад в данную область науки.

Опубликованные автором работы по теме диссертации хорошо отражают содержание диссертации. Автореферат правильно и полностью отражает содержание диссертации.

Диссертационная работа Арсена Аргашесовича Аракеляна по теме “Разработка биоинформатических подходов для изучения молекулярных механизмов развития хронических заболеваний человека” полностью соответствует требованиям ВАК РА, предъявляемым к докторским диссертациям, а автор работы заслуживает присуждения ему ученой степени доктора биологических наук по специальности 03.00.03 - "Молекулярная и клеточная биология".

Профессор кафедры молекулярной физики
Ереванского Государственного Университета,
доктор физико-математических наук

 В.Б. Аракелян

Подпись профессора В.Б. Аракеляна заверяю:

Ученый секретарь
Ереванского Государственного Университета,
кандидат исторических наук

 Л.С. Овсепян



29.11.2019