ПЕРСОНАЛИЗИРОВАННАЯ МЕДИЦИНА В ОДИН КЛИК: АВТОМАТИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ТЕРАПЕВТИЧЕСКОЙ МИШЕНИ ПО ДНК И РНК ПАЦИЕНТА

УДК 004.942, 004.67,
577.25, 577.29, 577.218
БАК 03.01.09

Стельмашенко Д.Е.^{1,2} (stelmaschenko@biosoft.ru), **Колпаков Ф.А.**, к.б.н.^{1,3}, **Евшин И.С.**¹, **Мандрик Н.В.**¹, **Рябова А.С.**¹, **Пинтус С.С.**¹, **Киселев И.Н.**, к.ф.-м.н.¹, **Шарипов Р.Н.**¹, **Штегмайер Ф. (Stegmaier P.)**², **Кошманн Д. (Koschmann J.)**, Ph.D.², **Кель-Маргулис О.В. (Kel-Margoulis O.V.)**, к.б.н.², **Крул М. (Krull M.)**², **Тандон Н. (Tandon N.)**², **Вингендер Э. (Wingender E.)**^{2,4}, **Кель А.Э. (Kel A.E.)**, к.б.н.²

Правильная идентификация терапевтической мишени – первый шаг на пути к успешному лечению таких сложных заболеваний, как онкологические новообразования. Новая компьютерная программа "Геномный усилитель" не только решает эту задачу за считанные часы в полностью автоматизированном режиме, но и осуществляет подбор оптимальной комбинации лекарственных препаратов, основанный на молекулярном механизме патологии конкретного пациента. "Геномный усилитель" позволяет вывести прежде глубоко научную задачу поиска терапевтических мишеней на практический уровень, доступный клиницистам и исследователям, не имеющим системно-биологических и биоинформационных компетенций.

Развитие геномных технологий и появление таргетной терапии открыли новую эру в медицине XXI века – эру персонализированного лечения пациентов, основанного на подавлении молекулярного механизма конкретной патологии. Особую роль персонализированная медицина и таргетные препараты играют в лечении онкологических заболеваний, поскольку в настоящий момент достоверно установлено, что рак одного и того же органа у двух разных людей может иметь совершенно отличную природу. Ввиду различий в молекулярных механизмах этих патологий некорректно использовать одно и то же лекарство для двух даже клинически идентичных заболеваний.

Для большинства раковых заболеваний были предложены различные классификации их подтипов, основанные на ассоциируемых с конкретной патологией генетических мутациях, метилировании или различных уровнях экспрессии конкретных генов. Все эти классификации помогают лечащим врачам так или иначе стратифицировать пациентов в отдельные молекулярные подгруппы. Пациенты одной подгруппы получают одинаковое лечение, однако зачастую ответ на про-

водимое лечение оказывается принципиально разным. Это связано не только с недостаточно корректной классификацией при стратификации пациентов, но и с вновь образующимися мутациями и изменением принципов регуляции раковых клеток. Проблема осложняется тем, что даже после первоначального успеха в лечении онкологического пациента у большинства больных возникают такие осложнения, как появление резистентности опухоли к применяемой терапии или активное метастазирование раковых образований в периферийные органы. Подбор лекарственных препаратов для онкологического пациента, столкнувшегося с этими осложнениями, как правило, проводится лечащим врачом на его собственное усмотрение на основании предыдущего опыта или информации из литературных источников, которая, за редким исключением, никак не коррелирует с изменениями молекулярного механизма патологии, произошедшими с раковыми клетками пациента.

Ключ к решению проблемы выбора оптимального лекарственного препарата лежит в правильном выборе терапевтической мишени. Идентификация

¹ ООО "БИОСОФТ.РУ", Новосибирск, Россия

² geneXplain GmbH, Wolfenbüttel, Germany

³ Конструкторско-технологический институт вычислительной техники СО РАН, Новосибирск, Россия

⁴ Institute of Bioinformatics, University Medical Center Göttingen (UMG), Göttingen, Germany

мишени по опосредованным признакам, таким как набор конкретных мутаций в образцах опухолевой ткани пациента или изменение уровня экспрессии или статуса метилирования конкретного гена, далеко не всегда позволяет провести оптимальную стратификацию пациентов. Единственный выход – это проведение полноценного анализа мультиомиксных данных (геномных, протеомных, транскриптомных, метаболомных и пр.) каждого пациента, что на настоящий момент является исключительно прерогативой научно-исследовательских работ и не используется в клинической практике ввиду крайне высокой трудоемкости и сложности анализа этих данных.

АВТОМАТИЗАЦИЯ ПОИСКА МИШЕНЕЙ

Одним из методов идентификации перспективных терапевтических мишеней является так называемый восходящий анализ [1–3], эффективность которого была показана в ряде крупных международных проектов, таких как SysCol, Mimomics, SysMedIBD, Resolve, Gerontoshield, Tempus, Exitox, Epitemab и пр. Сравнивая два состояния клетки, данный метод позволяет выявить мастер-регуляторы – гены и белки, которые запускают в патологических клетках каскад сигнальных реакций, приводящих к неправильной работе этих клеток. Поиск мастер-регуляторов осуществляется в два этапа. На первом этапе находят набор дифференциально экспрессирующихся генов (ДЭГ), отличающих здоровые клетки от патологических, после чего при помощи базы данных сайтов связывания транскрипционных факторов TRANSFAC определяют комплексы транскрипционных факторов, регулирующих данные ДЭГ [4–7]. Сайты посадки транскрипционных факторов (ТФ) ищутся в регуляторных районах ДЭГ путем сканирования промотерных регионов этих генов библиотекой позиционных весовых матриц базы данных TRANSFAC. На втором этапе при помощи базы данных внутриклеточных сигнальных и метаболческих путей TRANSPATH находят общие регуляторы предсказанных на первом шаге ТФ. Данные регуляторы, по сути, контролируют весь процесс передачи патологических сигналов внутри клетки, и именно они называются мастер-регуляторами. Мастер-регуляторы служат перспективными кандидатами в терапевтические мишени, поскольку их ингибирование приводит к блокировке сигналов, приводящих к активации дифференциально экспрессирующихся в патологии генов.

Для автоматизации процесса поиска терапевтических мишеней нами был использован и существенно доработан алгоритм восходящего анализа (рис.1).

Рис.1. Схема поиска мастер-регуляторов

Мы приспособили данный алгоритм для работы со многими типами омиксных данных (геномных, протеомных, транскриптомных и эпигеномных) и осуществили с его помощью интеграцию омиксных данных различных типов в единый алгоритм анализа. На вход алгоритму подается набор любых омиксных данных пациента, который напрямую зависит от проведенных ранее экспериментальных исследований образцов изучаемой патологической ткани или биологической жидкости больного. В зависимости от комплектности набора поданных на вход алгоритму данных, в анализ включаются те или иные блоки их обработки.

Так, для геномных данных всех типов (данные таргетного, полноэкзомного или полногеномного секвенирования, а также готовые .bam или .vcf файлы) разработанный алгоритм поиска терапевтических мишеней идентифицирует мутировавшие в изучаемом патологическом процессе гены и генерирует их список в качестве промежуточного набора генов-кандидатов для дальнейшего поиска регулирующих их ТФ. В случае если во входном наборе омиксных данных пациента имелись также данные генной экспрессии патологических и контрольных образцов (на основе полногеномного и частичного анализа транс-

криптома – данные микрочипов, RNA-seq или даже просто таблично-текстовые списки генов), для дальнейшего анализа и поиска ТФ будут отобраны те гены, которые одновременно являются мутировавшими и дифференциально экспрессирующимися в патологических образцах по сравнению с контролем. Такое пересечение имеет глубокий биологический смысл: для анализа отбираются реально работающие в патологических клетках мутировавшие гены. Аналогично, если во входном наборе омиксных данных пациента были эпигенетические данные по метилированию генома, то для дальнейшего анализа будут отобраны регуляторные районы генов, попадающие на участки открытого хроматина, поскольку именно эти районы с наибольшей вероятностью регулируют активность генов, участвующих в каскадах "неправильных" внутриклеточных сигналов.

После составления списка генов, оптимально описывающих изучаемую патологию, при помощи базы данных TRANSFAC осуществляется поиск комплексов ТФ, регулирующих данные гены [4–7]. Предсказанные наборы ТФ подаются на вход второму этапу метода восходящего анализа. На этом этапе при помощи базы данных TRANSPATH [8] осуществляется поиск ближайших общих регуляторов выявленных ТФ путем построения персонализированной сети передачи сигналов для изучаемой патологии. При поиске общих регуляторов выбранных ТФ, а именно сигнальных генов и белков, участвующих в трансдукции патологических сигналов, используется алгоритм поиска кратчайшего пути в графе. Наличие в исходном наборе экспериментальных транскриптомных или протеомных данных позволяет присваивать вершинам графа (генам и белкам) веса, что делает персонализированную сеть передачи сигнала более точной. При поиске мастер-регуляторов построенной сети учитывается не только роль этих генов и белков в общей системе сигнальной трансдукции патологии, но и возможность воздействия на данные молекулы лекарственными веществами (потенциал ингибирования или активирования выбранных мастер-регуляторов известными терапевтическими препаратами или классами малых молекул). При окончательном выборе наиболее перспективных мастер-регуляторов (терапевтических мишеней) учитывается также наличие образуемых данными мастер-регуляторами положительных обратных связей – "замыканий" системы, при которых гены нижнего уровня продуцируют белки, являющиеся ключевыми сигнальными молекулами системы передачи патологических сигналов клетки.

РЕШЕНИЕ НА ОСНОВЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Алгоритм автоматизированного поиска терапевтических мишеней был реализован в программном продукте "Геномный усилитель", который принимает на вход набор мультиомиксных данных пациентов и генерирует отчет, содержащий описание персонализированной модели молекулярного механизма изучаемой патологии. "Геномный усилитель" не только идентифицирует терапевтические мишени, но и подбирает перспективные лекарственные вещества, способные на них воздействовать. Поиск лекарственных препаратов осуществляется путем сканирования базы данных HumanPSD, содержащей информацию обо всех зарегистрированных лекарствах, а также лекарственных веществах, проходящих в настоящий момент клинические испытания. Такой мишень-центрированный подход к выбору перспективного лекарственного препарата предполагает возможность использования известных лекарств off-label, что все чаще становится повседневной клинической практикой в ведущих медицинских учреждениях мира.

Помимо подбора известных лекарственных препаратов, "Геномный усилитель" осуществляет также поиск перспективных низкомолекулярных соединений, способных ингибировать или активировать идентифицированные мишени. Подбор малых молекул производится путем сканирования заранее подготовленной при помощи программы PASS [9–11] библиотеки соединений с предсказанными биологическими активностями. Все включенные в библиотеку малые молекулы являются реальными уже синтезированными химическими веществами, которые можно приобрести у провайдеров. Это означает, что предсказанный "Геномным усилителем" механизм подавления патологии можно легко провалидировать экспериментально на клеточных линиях или модельном организме. Безусловно, предсказанные низкомолекулярные соединения нельзя напрямую давать пациентам, однако эта опция воздействия на молекулярный механизм изучаемой патологии будет крайне интересна клиническим лабораториям и научным организациям, занимающимся собственными исследованиями.

"Геномный усилитель" – крайне простой в использовании инструмент (рис.2). От пользователя программы требуется лишь загрузить в нее исходные экспериментальные данные и указать системе, какие из файлов относятся к изучаемому патологическому процессу, а какие – являются контролем. Так, например, при изучении солидных раковых опухолей омиксные данные, полученные из опухоли, будут служить образцами изучаемой патологии, а данные по здоровым клеткам того же органа пациента – контролем. После загрузки данных и разделения

их на образцы патологии и контроля пользователю остается только выбрать название изучаемого заболевания и тип ткани, из которой был осуществлен забор биологических образцов. Всего через несколько часов после нажатия кнопки запуска анализа пользователь получит готовый отчет, содержащий описание молекулярного механизма изучаемой патологии, а также списки идентифицированных терапевтических мишеней, терапевтических препаратов и лекарственных веществ, способных воздействовать на данные мишени.

"Геномный усилитель" не требует установки на локальный компьютер пользователя и предоставляет свой анализ в виде онлайн-сервиса, реализованного на облачных серверах.

ПРИМЕР ПРИМЕНЕНИЯ ТЕХНОЛОГИИ

В качестве примера применения "Геномного усилителя" нами были взяты RNA-seq-данные секвенирования опухолей колоректального рака пациентов из когорты, которая исследовалась в международном проекте SysCol (<http://syscol-project.eu/>). Полное исследование этих данных, включающих около 300 образцов, проводилось нами ранее и позволило идентифицировать набор биомаркеров метилирования ДНК, выявляемых в крови большинства этих пациентов. Данные маркеры позволяют проводить раннюю диагностику колоректального рака [3]. Однако, несмотря на наличие общих биомаркеров, пациенты когорты значительно различались как по общему профилю экспрессии генов, так и по течению болезни. Многие больные, имевшие изначально достаточно схожую клиническую картину заболевания, показывали принципиально разный ответ на проводимую адъювантную терапию и, как следствие, имели совершенно разные показатели продолжительности жизни с момента начала лечения. Таким образом, большой интерес представлял сравнительный анализ данных таких пациентов в целях выявления причин столь разного ответа на терапию.

Для анализа нами были выбраны данные двух пациентов, которые имели одинаковую стадию развития заболевания (вторая стадия онкологии), но в конечном счете – значительное отличие в продолжительности жизни после проведения лечения. Первый пациент жил 785 дней после операции (пациент ShortTerm), второй пациент – более 3275 дней (пациент LongTerm).

С помощью программы "Геномный усилитель" нами были проанализированы RNA-seq-данные этих пациентов, причем для каждого из них в анализ были включены отсековированный образец опухолевой ткани и образец здоровой ткани того же пациента, который использовался в качестве контроля. Секвенирование всех образцов проводилось по технологии парных концевых

Рис.2. Программа "Геномный усилитель": а) интерфейс загрузки данных; б) загруженные данные для проведения анализа; в) запуск анализа

Рис.3. Аннотация данных в "Геномном усилителе"

фрагментов, на вход "Геномному усилителю" подавалось по два fastq файла на каждый из образцов.

После загрузки данных была проведена простая аннотация: образцы были распределены по соответствующим группам патологии (Т) и контроля (N) для пациентов LongTerm (LongSurvival боксы на рис.3) и ShortTerm (ShortSurvival боксы на рис.3).

Далее для каждого из случаев (пациент LongTerm и пациент ShortTerm) в программе "Геномный усилитель" был запущен анализ, проводящий сравнение образцов нормальной и опухолевой тканей каждого из пациентов по отдельности. В результате анализа был сгенерирован отчет, содержащий описание персонализированного молекулярного механизма патологии каждого из пациентов. Помимо списков идентифицированных ДЭГ, позволяющих разделить случаи нормальной и патологической ткани и распределения данных генов по различным функциональным онтологиям и сигнальным путям, были указаны ключевые комплексы транскрипционных факторов, регулирующих выявленные ДЭГ.

На рис.4 приведен пример того, как в отчете "Геномного усилителя" визуализируются идентифицированные регуляторные комплексы ТФ, регулирующие ДЭГ и отличающие пациентов ShortTerm от пациентов LongTerm. Анализ выявил девять мотивов, или позиционных весовых матриц базы данных TRANSFAC, для различных ТФ, входящих в эти комплексы. Комплексы совместно работающих ТФ обозначены на рис.4 как модули 1 и 2, а мотивы – как синие прямоугольники внутри этих модулей. Совокупность данных модулей представляет собой композиционную модель, классифицирующую промоторы ДЭГ пациентов ShortTerm и LongTerm по сравнению с промоторами генов, имевших одинаковую экспрессию у обоих пациентов. На гистограмме внизу рис.4 показаны распределения показателей (scores), вычисленных с помощью этой композиционной модели для каждого из исследуемых промоторов генов. Красным цветом обозна-

чены показатели для промоторов ДЭГ, а синим цветом – для не изменивших свою экспрессию генов. Видно, что эти два распределения достоверно различаются (значение Z-score для AUC равно 1.6, что означает, что это значение достоверно отличается от значений AUC для случайных выборок промоторов).

Далее в отчете приводится диаграмма, визуализирующая персонализированную сигнальную сеть для случая патологии пациента ShortTerm (рис.5). Мастер-регуляторы отмечены красными прямоугольниками. Фиолетовые (предпоследняя строчка) и синие (последняя строчка) прямоугольники представляют собой идентифицированные на предыдущем шаге анализа ТФ. Зеленые прямоугольники – это промежуточные молекулы, участвующие в передаче патологических сигналов клетки. Синяя и красная обводки вокруг прямоугольников показывают даунрегуляцию и апрегуляцию соответствующих генов в молекулярном механизме патологии.

Для лечения заболевания с выявленным молекулярном механизмом "Геномный усилитель" предлагает список лекарственных веществ, которые, согласно прогнозу, должны оптимально подойти пациенту ShortTerm. Как видно из рис.5 и 6, программой было обнаружено большое

Рис.4. Выявленные комплексы ТФ для случая ShortTerm

**9–13 сентября 2019 года
Санкт-Петербург**

МЕНДЕЛЕЕВСКИЙ СЪЕЗД
по общей и прикладной химии

Уважаемые коллеги!

***Приглашаем вас принять участие в работе
XXI Менделеевского съезда по общей и прикладной химии,
который состоится с 9 по 13 сентября 2019 года в
Санкт-Петербурге и станет одним из основных мероприятий
Международного года Периодической таблицы химических
элементов, провозглашённого ООН в декабре 2017 г.***

В 2019 г. исполняется 150 лет создания Д.И. Менделеевым Периодической таблицы химических элементов – самого известного в мире научного достижения России за все времена существования отечественной науки.

Важно и то, что в 2018 г. исполнилось 150 лет Русского химического общества, созданного по инициативе Д.И. Менделеева при Санкт-Петербургском университете (сейчас – российское химическое общество имени Д.И. Менделеева).

Именно поэтому местом проведения XXI съезда выбран Санкт-Петербург – крупный центр химической науки и промышленности России. Тематика съезда охватывает основные направления развития химической науки, технологии и промышленности, химического образования и взаимодействия бизнеса с наукой и промышленностью, чем существенно отличаются от обычных тематических научных конференций.

В работе XXI Менделеевского съезда в Санкт-Петербурге планируется участие 2200–2500 человек, в том числе 300 иностранных учёных, включая лауреатов Нобелевской премии. Планируется также участие российских и зарубежных компаний, связанных с производством химических продуктов и материалов, в том числе фармацевтических субстанций (с презентацией результатов своей деятельности и выпускаемой продукции). Это, безусловно, будет способствовать привлечению внимания общественности к науке в целом.

Участие представителей различных научных организаций и высших учебных заведений в работе съезда будет способствовать укреплению научного сотрудничества в Российской Федерации, плодотворному обмену идеями между учёными и их вовлечению в совместные научно-исследовательские и образовательные проекты и программы.

Мы будем рады видеть вас в качестве участников съезда и деловых партнёров предприятий и организаций Санкт-Петербурга, а также других регионов России.

Ждём вас в Санкт-Петербурге!

www.mendeleev2019.ru

Оргкомитет XXI Менделеевского съезда

Рис.5. Диаграмма, визуализирующая молекулярный механизм патологии пациента ShortTerm

количество активированных киназ, которые в свою очередь активируют работу множественных сигнальных путей в опухолевых клетках пациента ShortTerm, что в конечном счете приводит к процессу их быстрого роста и деления. "Геномный усилитель" выбрал для лечения пациента ShortTerm известный препарат регорафениб (regorafenib) и другие лекарства, ингибирующие широкий спектр киназ (табл.1). Выбранные препараты характеризуются высокой степенью агрессивности (а также большим количеством побочных эффектов) и показаны для лечения агрессивно растущих опухолей колоректального рака.

Диаграмма выявленного "Геномным усилителем" молекулярного механизма заболевания для случая пациента с большим показателем продолжительности жизни (пациент LongTerm) приведена на рис.6.

При сравнении рис.5 и 6 можно заметить, что одно из главных отличий между выявленными сетями передачи сигналов данных пациентов состоит в полностью противоположной активности комплекса фактора роста EGF с его рецептором (EGF/ErbB1/ErbB2/Src). Если для пациента ShortTerm была показана высокая активность этого комплекса (на рис.5 этот комплекс показан с красной подсветкой), то для пациента LongTerm наблюдалась пониженная активность этого комплекса (на рис.6 этот комплекс показан с синей подсветкой). По нашему мнению, именно высокая активность EGF комплекса – один из ключевых факторов быстрой прогрессии болезни у пациента ShortTerm.

Лекарственные препараты, выбранные "Геномным усилителем" для пациента LongTerm, менее агрессивны по срав-

Табл.1. Оптимальная комбинация лекарственных препаратов для пациента ShortTerm

Disease activity score	11	6	4
Phase 4	Colorectal Neoplasms, Gastrointestinal Stromal Tumors, Neoplasms, Rectal Neoplasms	Carcinoma, Renal Cell, Gastrointestinal Neoplasms, Gastrointestinal Stromal Tumors, Intestinal Neoplasms, Lung Neoplasms, Neoplasms, Neuroendocrine Tumors...	Carcinoma, Hepatocellular, Carcinoma, Renal Cell, Liver Neoplasms, Neoplasms, Noma, Thrombosis
Phase 3	Colorectal Neoplasms, Carcinoma, Hepatocellular, Colonic Neoplasms, Esophageal Neoplasms, Gastrointestinal Stromal Tumors, Neoplasms, Noma...	Colorectal Neoplasms, Brain Neoplasms, Breast Neoplasms, Carcinoma, Carcinoma, Hepatocellular, Carcinoma, Islet Cell, Carcinoma, Non-Small-Cell Lung...	Adenocarcinoma, Breast Neoplasms, Carcinoma, Carcinoma, Hepatocellular, Carcinoma, Non-Small-Cell Lung, Carcinoma, Renal Cell, Digestive System Diseases...
Phase 2	Colorectal Neoplasms, Adenocarcinoma, Bile Duct Neoplasms, Brain Abscess, Breast Neoplasms, Carcinoid Tumor, Carcinoma, Adenoid Cystic...	Colorectal Neoplasms, Acquired Immunodeficiency Syndrome, Adenocarcinoma, Clear Cell, Adenoma, Islet Cell, Adrenocortical Carcinoma, Ascites...	Colorectal Neoplasms, Adenocarcinoma, Adenoma, Adenoma, Liver Cell, Adrenocortical Carcinoma, Bile Duct Neoplasms, Biliary Tract Neoplasms...
Phase 1	Colorectal Neoplasms, Carcinoma, Hepatocellular, Carcinoma, Small Cell, Esophageal Neoplasms, Gastrointestinal Neoplasms, Gastrointestinal Stromal Tumors, Intestinal Neoplasms...	Colorectal Neoplasms, Acquired Immunodeficiency Syndrome, Adenocarcinoma, Amyloidosis, Anger, Brain Abscess, Brain Neoplasms...	Colorectal Neoplasms, Adenocarcinoma, Adenoma, Adenoma, Liver Cell, Astrocytoma, Bile Duct Neoplasms, Biliary Tract Neoplasms...
NA	Colorectal Neoplasms, Adenocarcinoma, Carcinoma, Hepatocellular, Cholangiocarcinoma, Gastrointestinal Stromal Tumors, Glioblastoma, Liver Neoplasms...	Adenocarcinoma, Breast Neoplasms, Carcinoid Tumor, Carcinoma, Hepatocellular, Carcinoma, Islet Cell, Carcinoma, Non-Small-Cell Lung, Carcinoma, Renal Cell...	Colorectal Neoplasms, Adenocarcinoma, Ascites, Brain Abscess, Brain Neoplasms, Breast Neoplasms, Carcinoma, Hepatocellular...
Target activity score	3.45	1.91	2.69
Target names	KIT, KDR, ABL1, PDGFRB, FGFR1, RET, PDGFRA...	KIT, KDR, PDGFRB, PDGFRA	KIT, KDR, PDGFRB, FGFR1, BRAF, RAF1, RET
Name	Regorafenib	Sunitinib	Sorafenib
ID	DB08896	DB01268	DB00398

нению с лекарствами, выбранными для пациента ShortTerm (табл.2). Среди них даже обычный витамин Е.

Таким образом, принципиальные отличия молекулярных механизмов патологий обуславливают раз-

личное оптимальное терапевтическое лечение для пациентов ShortTerm и LongTerm. К сожалению, данные о молекулярном механизме развития опухолей этих двух пациентов на момент принятия решения

Рис.6. Диаграмма, визуализирующая молекулярный механизм патологии пациента LongTerm

Табл.2. Оптимальная комбинация лекарственных препаратов для пациента LongTerm

Disease activity score	3	3	3	2	7
Phase 4	Angina Pectoris, Variant, Asphyxia, Cicatrix, Cicatrix, Hypertrophic, Diabetes Mellitus, ...	Leukemia, Leukemia, Lymphoid, Leukemia, Myelogenous, Chronic, BCR-ABL Positive, Leukemia, Myeloid...	Leukemia, Myeloid	Acne Vulgaris, Affect, Alopecia, Autistic Disorder, Bacterial Infections, Bipolar Disorder, ...	Acute Coronary Syndrome, Atherosclerosis, Atrial Fibrillation, Brain Abscess, Cardiovascular ...
Phase 3	Colorectal Neoplasms, Alzheimer Disease, Angina, Unstable, Arterial Occlusive Diseases, ...	Leukemia, Leukemia, Lymphoid, Leukemia, Myelogenous, Chronic, BCR-ABL Positive, Leukemia, Myeloid...	Leukemia, Leukemia, Myelogenous, Chronic, BCR-ABL Positive, Leukemia, Myeloid	Acne Vulgaris, Affect, Alopecia, Amphetamine-Related Disorders, Amyotrophic Lateral Sclerosis, ...	Colorectal Neoplasms, Acute Coronary Syndrome, Affect, Aneurysm, Atrial Fibrillation, ...
Phase 2	Acute Kidney Injury, Adrenoleukodystrophy, Anemia, Arsenic Poisoning, Arteriosclerosis, ...	Colorectal Neoplasms, Adenocarcinoma, Adenocarcinoma, Clear Cell, Blast Crisis, Brain Abscess, ...	Colorectal Neoplasms, Brain Abscess, Breast Neoplasms, Cholangiocarcinoma, Cysts, Glioblastoma, ...	Colorectal Neoplasms, Acne Vulgaris, Alcohol Drinking, Alopecia, Alzheimer Disease, Amyotrophic ...	Colorectal Neoplasms, Acute Coronary Syndrome, Arteriosclerosis, Atrial Fibrillation, Barrett ...
Phase 1	Alzheimer Disease, Anemia, Anemia, Iron-Deficiency, Arthritis, Arthritis, Rheumatoid, Asthma, ...	Colorectal Neoplasms, Adenocarcinoma, Adenocarcinoma, Clear Cell, Adenocarcinoma, Mucinous, Brain...	Colorectal Neoplasms, Acute Kidney Injury, Breast Neoplasms, Carcinoma, Non-Small-Cell Lung, ...	Acne Vulgaris, Acute Kidney Injury, Affect, Alcohol Drinking, Alcoholism, Alopecia, Aneurysm, ...	Colorectal Neoplasms, Arteriosclerosis, Drug-Related Side Effects and Adverse Reactions, Fatigue,...
NA	Apnea, Fatty Liver, Fatty Liver, Alcoholic, Fragile X Syndrome, HIV Infections, Hepatitis, ...	Brain Neoplasms, Carcinoma, Squamous Cell, Carcinoma, Transitional Cell, Gastrointestinal Stromal...	Breast Neoplasms, Leukemia, Leukemia, Lymphoid, Leukemia, Myelogenous, Chronic, BCR-ABL Positive,...	Acne Vulgaris, Acute Kidney Injury, Alopecia, Angelman Syndrome, Anxiety Disorders, ...	Colorectal Neoplasms, Abortion, Spontaneous, Adenocarcinoma, Affect, Arthritis, Arthritis, ...
Status (provided by Drugbank)	Small molecule, approved, nutraceutical	small molecule, approved, investigational	small molecule, approved	small molecule, approved, investigational	small molecule, approved
Target activity score	6.49108	5.97148	5.95753	9.18532	4.14619
Target names	PRKCA, PPP2CA	YES1, FYN	MAP2K1, LYN, CDK2	VEGFA, IL1B, CASP3, CYCS	PRKAA1, TP53
Name	Vitamin E	Dasatinib	Bosutinib	Minocycline	Acetylsalicylic acid
ID	DB00163	DB01254	DB06616	DB01017	DB00945

об их лечении отсутствовали, и лечащий врач назначил обоим пациентам одинаковое лечение, что привело к крайне неблагоприятному исходу для пациента ShortTerm. Перспектива действительно персонализированного подхода к лечению заболеваний, которую открывает перед пациентами и врачами "Геномный усилитель", могла бы существенно улучшить качество и продолжительность жизни многих онкологических пациентов.

ВЫВОДЫ

"Геномный усилитель" выводит на принципиально новый уровень наше представление о персонализированной медицине в целом и комплексном биоинформационном анализе в частности. До настоящего момента понятие персонализированной медицины включало методы оптимальной стратификации пациентов по заранее распределенным терапевтическим группам в соответствии с определенными биомаркерами, наблюдающимися в патологии. С приходом в клиническую практику "Геномного усилителя" у врачей появится возможность не решать, к какой из заранее сформированных групп следует отнести данного пациента, а выявлять реально активированные в текущий момент времени терапевтические мишени, включенные в патологический процесс

в пораженных клетках пациента. Лечить болезнь здесь и сейчас – вот ключевой девиз "Геномного усилителя".

Генерируемые "Геномным усилителем" результаты немаловажны и для научно-исследовательских организаций и лабораторий, проводящих обработку омиксных данных и не имеющих собственных высококвалифицированных кадров в области биоинформатики и системной биологии. Полная автоматизация процесса анализа с генерированием подробного отчета о проведенном исследовании позволяет биологам, врачам, лаборантам и техникам самостоятельно проводить комплексный биоинформационный анализ, который занял бы долгие месяцы в случае ручного его выполнения даже самыми квалифицированными кадрами.

"Геномный усилитель" расширяет наше представление о доступности методов обработки сырых омиксных данных и глубине и качестве результатов, которые могут быть сгенерированы искусственным интеллектом лишь по нажатию одной кнопки запуска системы.

Авторы благодарят Фонд развития Центра разработки и коммерциализации новых технологий (Фонд "Сколково") за поддержку реализации проекта "Геномный усилитель".

ЛИТЕРАТУРА

1. Koschmann J., Bhar A., Stegmaier P., Kel A.E., Wingender E. "Upstream Analysis": An integrated promoter-pathway analysis approach to causal interpretation of microarray data // Microarrays. 2015. V. 4. P. 270–286.
2. Kel A.E., Stegmaier P., Valeev T., Koschmann J., Poroikov V., Kel-Margoulis O.V., Wingender E. Multi-omics "upstream analysis" of regulatory genomic regions helps identifying targets against methotrexate resistance of colon cancer // EuPA Open Proteomics. 2016. V. 13. P. 1–13.
3. Kel A., Boyarskikh U., Stegmaier P., Leskov L.S., Sokolov A.V., Yevshin I., Mandrik N., Stelmashenko D., et.al. Walking pathways with positive feedback loops reveal DNA methylation biomarkers of colorectal cancer // BMC Bioinformatics. 2019. V. 20 (Suppl 4).
4. Kel A., Voss N., Jauregui R., Kel-Margoulis O., Wingender E. Beyond microarrays: find key transcription factors controlling signal transduction pathways // BMC Bioinformatics. 2006. V. 7 (Suppl 2). P. S13.
5. Matys V., Kel-Margoulis O.V., Fricke E., Liebich I., Land S., et.al. TRANSFAC and its module TRANSCOMP: transcriptional gene regulation in eukaryotes // Nucleic Acids Res. 2006. V. 34. P. D108–D110.
6. Kel A.E., Gössling E., Reuter I., Cheremushkin E., Kel-Margoulis O.V., Wingender E. MATCH: A tool for searching transcription factor binding sites in DNA sequences // Nucleic Acids Res. 2003. V. 31. P. 3576–3579.
7. Waleev T., Shtokalo D., Konovalova T., Voss N., Cheremushkin E., et.al. Composite Module Analyst: identification of transcription factor binding site combinations using genetic algorithm // Nucleic Acids Res. 2006. V. 34. P. W541–5.
8. Krull M., Pistor S., Voss N., Kel A., et.al. TRANSPATH: an information resource for storing and visualizing signaling pathways and their pathological aberrations // Nucleic Acids Res. 2006. V. 34. P. D546–D551.
9. Filimonov D.A., Poroikov V.V. Probabilistic approach in activity prediction. In: Chemoinformatics Approaches to Virtual Screening. // Eds. A. Varnek and A. Tropsha. Cambridge (UK): RSC Publishing. 2008. P. 182–216.
10. Filimonov D.A., Poroikov V.V. Prognosis of spectra of biological activity of organic molecules // Russian chemical journal. 2006. V. 50, № 2. P. 66–75.
11. Filimonov D., Poroikov V., Borodina Yu., Glorizova T. Chemical Similarity Assessment through Multilevel Neighborhoods of Atoms: Definition and Comparison with the Other Descriptors // J. Chem. Inf. Comput. Sci. 1999. V. 39. P. 666–670.

Валидация аналитических методик. Неопределенность в аналитических измерениях.
Под ред. Г.Р.Нежиховского, Р.Л.Кадиса

ИСО/МЭК 17025:2017. Практические рекомендации по применению.
И. В. Болдырев

Анализ почвы. Справочник.
Пансю М., Готеру Ж.

Анализ масел. Основы и применение.
Дж. Фитч, Д. Тройер

Индуктивно-связанная плазма. Практическое руководство.
Дж.Дин

Химия окружающей среды.
Стэнли Е. Монаган

Контроль качества в аналитической химии.
Причард Э., Барвик В.

Производство лекарственных средств. Контроль качества и регулирование.
Ш.К.Гэд

Основы реологии.
А.Я. Малкин