

УДК 004.942; 001.5; 577.1; 577.2; 616.03

А.И. Глухов

д-р биол. наук, профессор, заведующий кафедрой биологической химии ФГАОУ ВО Первый МГМУ им. И.М. Сеченова (Сеченовский Университет); профессор биологического факультета ФГБОУ ВО «Московский государственный университет им. М.В. Ломоносова»

З.А. Хучуа

канд. биол. наук, доцент кафедры биологической химии ФГАОУ ВО Первый МГМУ им. И.М. Сеченова (Сеченовский Университет); ассоциированный профессор Детского медицинского исследовательского центра, г. Цинциннати, Огайо, США

Г.К. Грызунова

канд. биол. наук, доцент кафедры биологической химии ФГАОУ ВО Первый МГМУ им. И.М. Сеченова (Сеченовский Университет)

Д.В. Астахов

канд. биол. наук, доцент кафедры биологической химии ФГАОУ ВО Первый МГМУ им. И.М. Сеченова (Сеченовский Университет)

М.И. Данилевский

канд. биол. наук, доцент кафедры биологической химии ФГАОУ ВО Первый МГМУ им. И.М. Сеченова (Сеченовский Университет)

A.I. Glukhov

PhD, DSci, Prof., Head of the Department of Biological Chemistry, I.M. Sechenov First Moscow State Medical University; Prof. of Biological Faculty, Lomonosov Moscow State University

Z.A. Khuchua

PhD, Associate Prof., Department of Biological Chemistry, I.M. Sechenov First Moscow State Medical University; Associate Prof., Cincinnati Children's Hospital and Medical Center, Cincinnati, Ohio, USA

G.K. Grizunova

PhD, Associate Prof., Department of Biological Chemistry, I.M. Sechenov First Moscow State Medical University

D.V. Astakhov

PhD, Associate Prof., Department of Biological Chemistry, I.M. Sechenov First Moscow State Medical University

M.I. Danilevskiy

PhD, Associate Prof., Department of Biological Chemistry, I.M. Sechenov First Moscow State Medical University

ИНТЕРАКТОМИКА В ТРАНСЛЯЦИОННОЙ МЕДИЦИНЕ INTERACTOMICS IN TRANSLATIONAL MEDICINE

КОНТАКТНАЯ ИНФОРМАЦИЯ

Глухов Александр Иванович, д-р биол. наук, профессор, заведующий кафедрой биологической химии ФГАОУ ВО Первый МГМУ им. И.М. Сеченова (Сеченовский Университет); профессор биологического факультета ФГБОУ ВО «Московский государственный университет им. М.В. Ломоносова»

Адрес: 119991, г. Москва, ул. Большая Пироговская, д. 2, стр. 4

Тел.: +7 (985) 295-70-05

E-mail: glukhov.ai@lmsmu.ru

Статья поступила в редакцию: 28.11.2017

Статья принята к печати: 15.02.2018

CONTACT INFORMATION

Glukhov Alexander, PhD, DSci, Prof., Head of the Department of Biological Chemistry, I.M. Sechenov First Moscow State Medical University; Prof. of Biological Faculty, Lomonosov Moscow State University

Address: 2–4, Bolshaya Pirogovskaya st., Moscow, 119991, Russia

Tel.: +7 (985) 295-70-05

E-mail: glukhov.ai@lmsmu.ru

Article received: November 28, 2017

Article approved for publication: February 15, 2018

Abstract

The article describes the place of interactomics in translational medicine. The approaches and methods used for the detection of protein-protein interactions and the construction of interactomes are described. Examples of electronic databases that work with interactomes are given, as well as examples of consortia of such databases. Examples of the use of interactomes for developing

примеры использования интерактомов для разработки новых подходов и методов лечения таких заболеваний, как рак, диабет, шизофрения. Обозначены основные трудности, возникающие при построении интерактомов, и перспективы развития интерактомики.

Ключевые слова: биоинформатика, интерактомика, трансляционная медицина.

new approaches and methods of treatment of such diseases as cancer, diabetes, schizophrenia are considered. The main difficulties encountered in the construction of interactomes, and the prospects for the development of interactomics are outlined.

Keywords: bioinformatics, interactomics, translational medicine.

Трансляционная медицина — это междисциплинарная область знаний, основной задачей которой является объединение активов фундаментальной биомедицинской науки с клинической медициной для совершенствования лечения и реабилитации пациентов, а также профилактики и диагностики заболеваний [1].

Научные прорывы, в том числе в геномике, биологии стволовых клеток и молекулярной биологии, за последние полвека обеспечили беспрецедентный доступ к информации, позволяющей повысить качество и продолжительность жизни человека. Трансляция, т. е. перевод этой информации, в знания, которые будут оказывать влияние на клиническую практику и в конечном счете на здоровье человека, потребует трансляционных исследований с участием людей.

Трансляционное исследование делится на пять отдельных фаз (T0–T4), известных как T-фазы, и включает в себя внедрение знаний, полученных из фундаментальной науки, в здравоохранение и профилактику заболеваний (рис. 1).

Первая фаза (T0) представляет собой основной этап фундаментальных исследований [2]. Лабораторные исследования, накопленные знания и научные открытия на фазе T0 являются базисом для проведения доклинических исследований и для подготовки клинических испытаний на людях на стадиях T1 и T2. Фазы T3 и T4 включают в себя перевод исследований клинических испытаний к усовершенствованию медицинской практики. Однако для того, чтобы трансляционное исследование действительно достигло этой конечной цели, оно должно иметь обратную связь с фундаментальной наукой (рис. 2).

Таким образом, с одной стороны, трансляция фундаментальных знаний в новые методы лечения и понимание механизмов болезней требуют проведения клинических испытаний, которые в конечном итоге помогут генерировать новые идеи, с другой стороны, клинические наблюдения за проявлениями болезни и результатами лечения должны, в свою очередь, эффективно доводиться до сведения ученых-исследователей, тем самым обеспечивая стимул для дальнейших фундаментальных исследований (см. рис. 1). Эта двойная направленность лежит в основе концепции «из лаборатории к больничной койке и обратно в лабораторию» (lab bench to the bedside and back to the bench). Следовательно, исследования в области трансляционной медицины несомненно увеличат потребность в финансировании как фундаментальной науки, так и клинических исследований.

Рынок глобальной трансляционной регенеративной медицины в 2016 г. оценивался в 5,8 млрд долларов США. Прогнозируется, что этот рынок увеличится до 14,5 млрд долларов США в 2021 г. [3].

Важным стимулирующим шагом для преодоления трансляционных барьеров является развитие трансдисциплинарного подхода с целью интеграции фундаментальных исследований в клиническую практику, чему способствует применение методов компьютерной обработки биомедицинских данных, т. е. биоинформатики.

Омиксы (omics) представляют собой данные, полученные в результате исследований с использованием высокопроизводительных технологий (секвенирование нового поколения, микрочи-

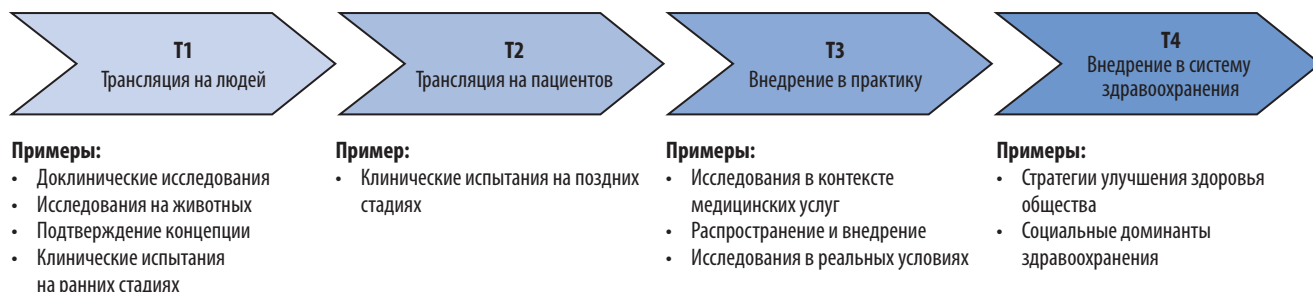


РИС. 1. Четыре фазы трансляционного процесса (схема, предложенная Институтом клинической и трансляционной медицины Калифорнийского университета в г. Лос-Анжелесе (США), <https://ctsi.ucla.edu/about/pages/featured>). Фаза T0 не указана.

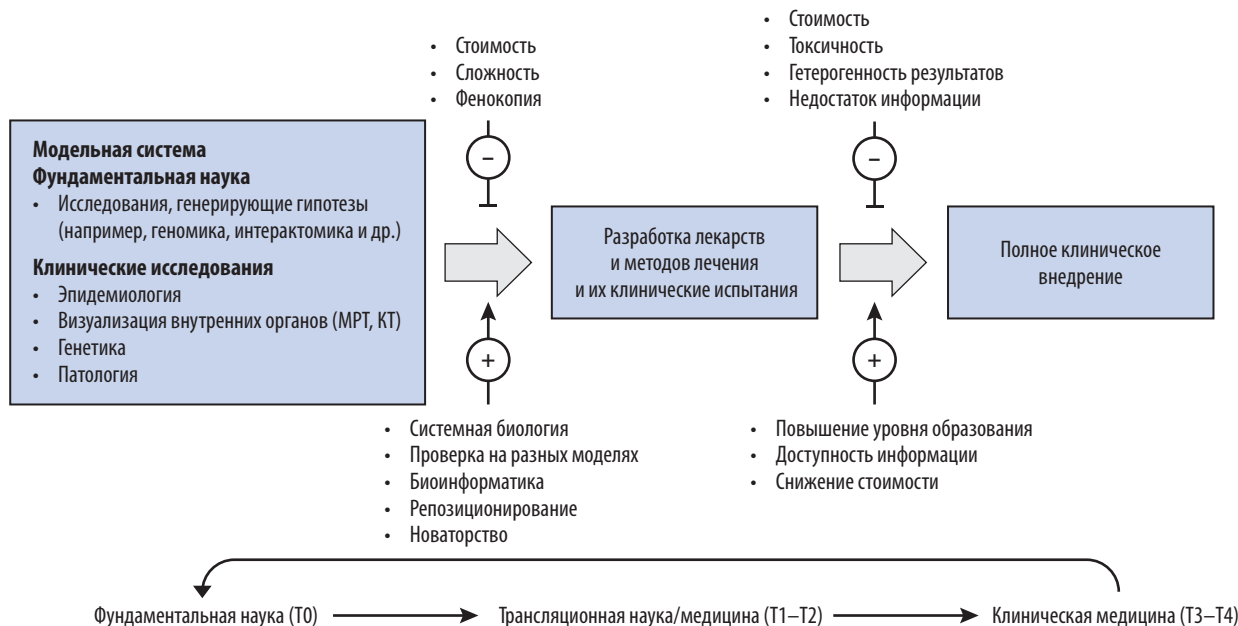


РИС. 2. Взаимодействие участников процесса трансляции фундаментальных знаний в клиническую медицину. Схема иллюстрирует поток информации и открытий от фундаментальной науки (слева) через трансляционную науку и медицину (в середине) до окончательного внедрения в клиническую практику (справа). Вертикальные линии отображают препятствующие (–) и стимулирующие (+) факторы, влияющие на трансляционный процесс. (Цит. по [2], с изм.)

пы, протеомика, тандемная масс-спектрометрия) на уровнях геномов, транскриптомов, протеомов, метаболомов, интерактомов и эпигеномов. Такие массивы данных можно обрабатывать только с использованием биоинформатических методов. Биоинформатика позволяет описывать и количественно просчитывать сложные биологические системы и клеточные сети, которые обуславливают соотношения генотипов и фенотипов. В 1941 г. Бидл и Татум [4] постулировали концепцию «один ген — один фермент — одна функция», которая содержит простые линейные связи между генотипом и фенотипом. Однако в действительности большинство отношений генотип–фенотип является результатом гораздо более сложных взаимодействий. Даже при анализе простых менделевских признаков наследования наблюдаются сложные отношения между генотипом и фенотипом. Неполная пенетрантность, переменная экспрессия, различия в возрасте при проявлении болезней встречаются чаще, чем оценивалось раньше. Однояйцевые близнецы, хотя и во многом сходны, тем не менее часто проявляют множество различий.

Взаимодействуя друг с другом, гены и их продукты образуют сложные сети взаимодействий внутри клеток — интерактом. Интерактом — это карта биологически значимых молекулярных взаимодействий. В настоящее время широко признается, что связь между интерактомом и фенотипом столь же важна, как связь между фенотипом и генотипом.

Молекулярные взаимодействия могут происходить между молекулами, принадлежащими к различным биохимическим семействам (белки, нуклеиновые кислоты, липиды, углеводы), а также между молекулами внутри отдельного семейства. Если такие молекулы связаны физическими взаимодействиями, они образуют сети молекулярных взаимодействий (интерактом), которые, как правило, классифицируются по характеру взаимодействующих молекул. Чаще всего к интерактому относятся сети белок-белковых взаимодействий (ББВ). Взаимодействующие компоненты (в данном случае белки) называются узлами (nodes), а взаимодействия между узлами отображаются линиями и называются ребрами (edges). Узлы, к которым сходится много ребер (рис. 3, желтые овалы), называются хабами (hubs). Математическая оценка предсказывает, что в интерактоме человека в целом существует 650 000 взаимодействий (ребер) между 25 000 белками (узлами) [5].

Другим широко изученным типом интерактома является взаимодействие белков с ДНК, также называемое ген-регуляторной сетью, или сетью, образованной факторами транскрипции, регуляторными белками хроматина и их генами-мишенями. Даже метаболические сети можно рассматривать как сети молекулярного взаимодействия, в которых метаболиты, т. е. химические соединения в клетке, преобразуются друг в друга с помощью ферментов, физически связывающих свои субстраты (рис. 4).

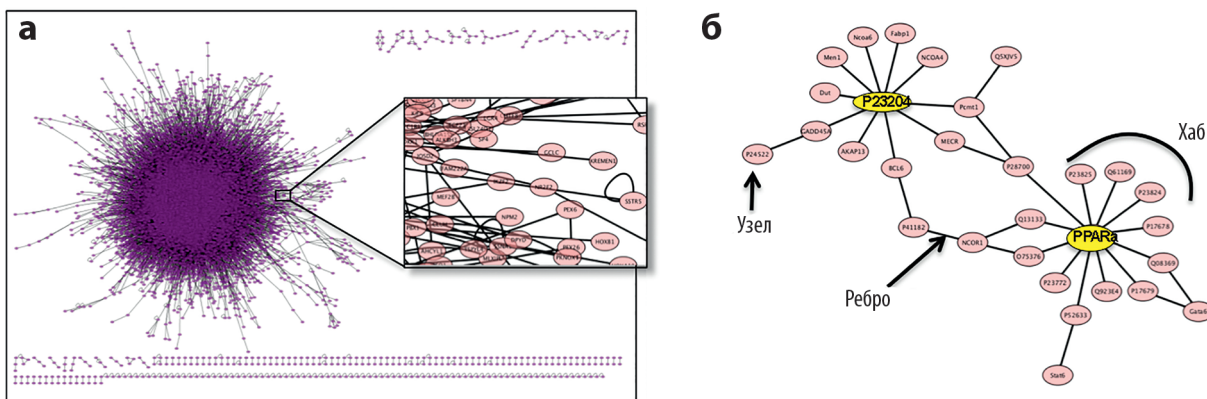


РИС. 3. Примеры белковых интерактомов: (а) полный интерактом человека, полученный из сервера APID (55 709 взаимодействий между 10 345 белками). Все взаимодействия в данном интерактоме подтверждены по крайней мере в двух публикациях; (б) интерактом ядерного рецептора PPAR α . Построен с помощью программного обеспечения Cytoscape 3.5

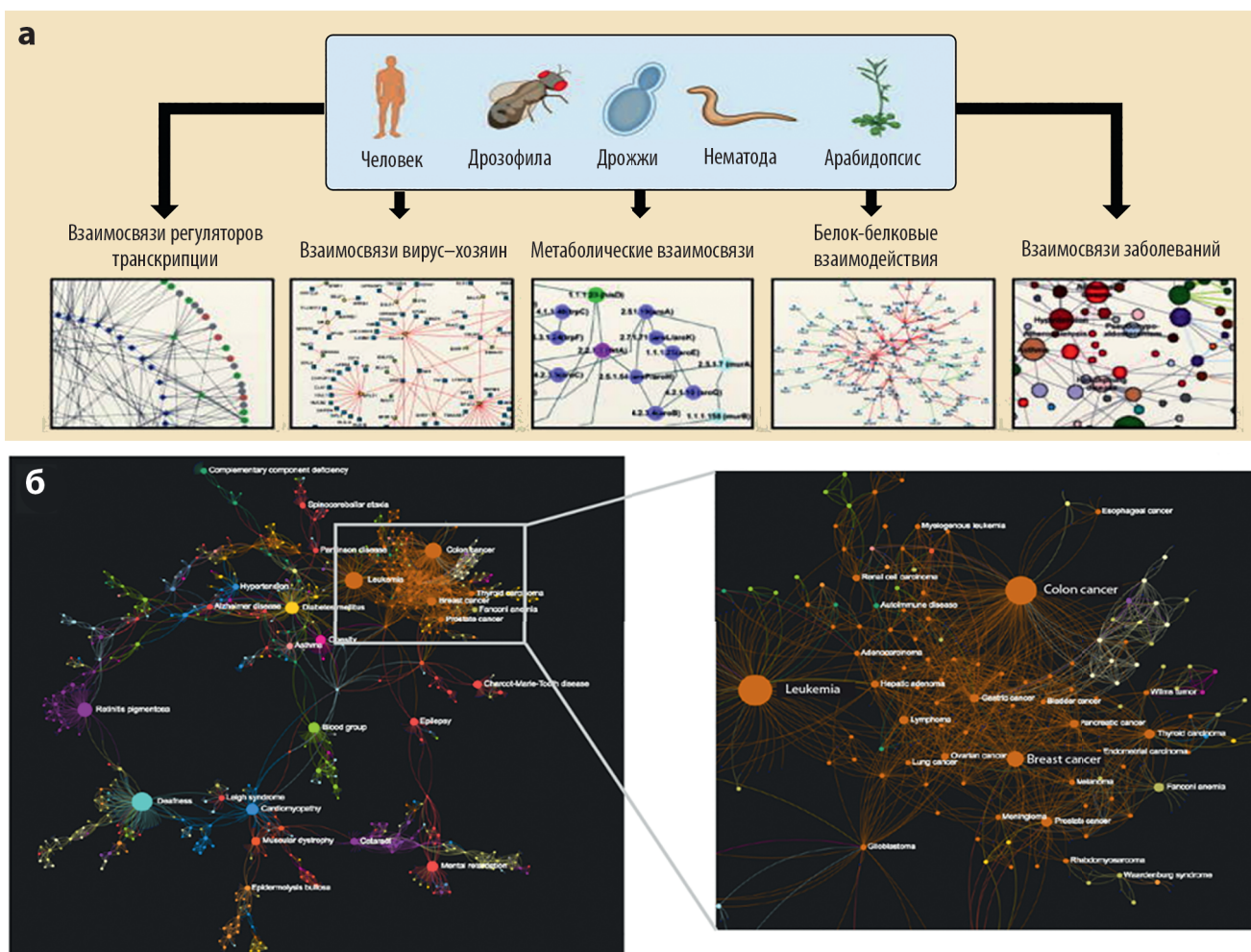


РИС. 4. Интерактомные сети в клеточных системах: (а) для «супермодельных» организмов человека, дрозофилы, дрожжей, червей и растений; в транскрипционных регуляторных сетях узлы представляют факторы транскрипции или предполагаемые регуляторные элементы ДНК, ребра — физическое взаимодействие между ними (в сети вирус–хозяин узлы — вирусные белки или белки хозяина, ребра — физические взаимодействия; в метаболической сети узлы — ферменты, ребра — метаболиты, которые являются продуктами или субстратами ферментов; в сети межбелковых взаимодействий узлы — белки, ребра — физические взаимодействия) (по Vidal M. et al [7]); (б) для генетических заболеваний: узлы представляют собой болезни, ребра — гены, мутации в которых вызывают заболевания. Визуализация произведена с помощью графического пакета Gephi (<https://gephi.org/>)

Одно из основных предположений, вытекающих из гипотезы о том, что болезни следует рассматривать как нарушение тесно взаимосвязанных клеточных сетей, заключается в том, что болезни не должны быть независимы друг от друга, а должны быть в высокой степени взаимосвязаны. Эта концепция привела к созданию различных карт глобальной сети болезней, в которых взаимосвязь фенотипов различных болезней имеет место только в том случае, если между ними существуют молекулярные или фенотипические связи. Диабет и ожирение представляют собой, вероятно, самую известную пару заболеваний с такими связями. Такая карта была построена с использованием известных ассоциаций генов и заболеваний, собранных в базе данных OMIM, где узлы являются заболеваниями, а два заболевания соединены ребром, если они имеют по меньшей мере один общий ген, мутации в котором ассоциированы с этими заболеваниями [6]. В полученной сети более 500 человеческих генетических болезней оказались относящимися к одному взаимосвязанному основному гигантскому кластеру (см. рис. 4, б). Это может свидетельствовать о том, что болезни человека гораздо более взаимосвязаны друг с другом, чем мы предполагаем.

На самом деле, все типы взаимодействий также связаны между собой. Например, интерактомы белковых взаимодействий содержат много ферментов, которые, в свою очередь, образуют биохимические сети. Аналогично сети регуляторных генов существенно перекрываются с сетями белковых взаимодействий и сигнальными сетями.

Основной элемент белкового интерактома — белок-белковое взаимодействие. В то время как существует достаточно много методов исследования белок-белковых взаимодействий, немногие из них могут использоваться при создании интерактомов. Для определения прямых физических взаимодействий между двумя белками используется двугибридная дрожжевая система (yeast two-hybrid system, Y2H). Метод достаточно прост, однако часто приводит к ложноположительным результатам. Вторым, более надежным, методом является аффинная очистка, например, ко-иммунопреципитация (co-IP), сопряженная с последующей масс-спектрометрией (MS). Плюсы этой технологии в том, что во время опыта исследуемые белки находятся в нативной конформации. Однако ко-иммунопреципитация может не уловить слабое взаимодействие и не может отличить прямое взаимодействие от непрямого (взаимодействие при участии белков-посредников). Другой подход, например метод комплементации белка (PCA), также успешно может применяться на практике. Более полный обзор методов обнаружения межбелковых взаимодействий изложен в работе Snider J. et al. [8].

Оба метода (co-IP и Y2H) имеют высокую производительность. Комбинация «ко-иммунопре-

ципитация + масс-спектрометрия»(co-IP + MS) не имеет таких проблем с ложноположительными и ложноотрицательными срабатываниями, как Y2H, и используется в качестве золотого стандарта. Но двугибридная дрожжевая система лучше показывает неспецифические физические взаимодействия, в то время как система co-IP + MS лучше демонстрирует функциональные белок-белковые взаимодействия *in vivo*.

Существует множество электронных баз интерактивных данных. Некоторые репозитории белок-белковых взаимодействий только фиксируют взаимодействия из литературы и делают их доступными для научного сообщества в структурированном виде. Другие фокусируются на предсказанных взаимодействиях или на картировании экспериментально определенных взаимодействий в одном отдельном организме. На сентябрь 2013 г. существовало более 125 различных баз данных по белок-белковым взаимодействиям [9]. Большинство из них независимо управляются и финансируются. Эти репозитории не координируют свою практику сбора и поддержки данных и не прикладывают усилий по предоставлению данных для научного сообщества. Часто сложно узнать точный характер взаимодействий, качество и источники исходных данных. Еще один уровень сложности появляется при использовании репозиториями различных условий стандартизации в описании данных о взаимодействии между белками, полученными из разных организмов. Кроме того, на практике некоторые базы данных контролируют только часть информации, попадающей в конкретную область их интересов, тогда как другие систематически изучают все имеющиеся данные.

Чтобы решить эти проблемы, несколько самых известных баз данных — по белковым взаимодействиям (DIP) [10], IntAct [11], по молекулярным взаимодействиям (MINT) [12], MatrixDB [13], по микробным белкам (MPIDB) [14], InnateDB [15], Interologous Interaction Database (I2D) [16] и Molecular Connections (<http://www.molecularconnections.com>) — создали консорциум «Международный молекулярный обмен» (IMEX) [17]. Это позволило собрать уникальный набор межбелковых взаимодействий, доступный с одного веб-портала (<http://www.imexconsortium.org>), и разработать общие стандарты контроля, терминологии и представления данных. Видеоресурс, иллюстрирующий работу с данными на портале IMEX, находится по адресу: <https://www.youtube.com/embed/sSSZUcm8PRM>.

Другая интегрированная и периодически обновляемая интерактивная база находится на сервере APID (Agile Protein Interactomes Data), (рис. 5 и 6). Сервер APID был построен с использованием стратегии, ориентированной на протеом, базы данных UniProt (<http://www.uniprot.org/proteomes>), а также Swiss-Prot и TrEMBL в качестве основных источников по иден-

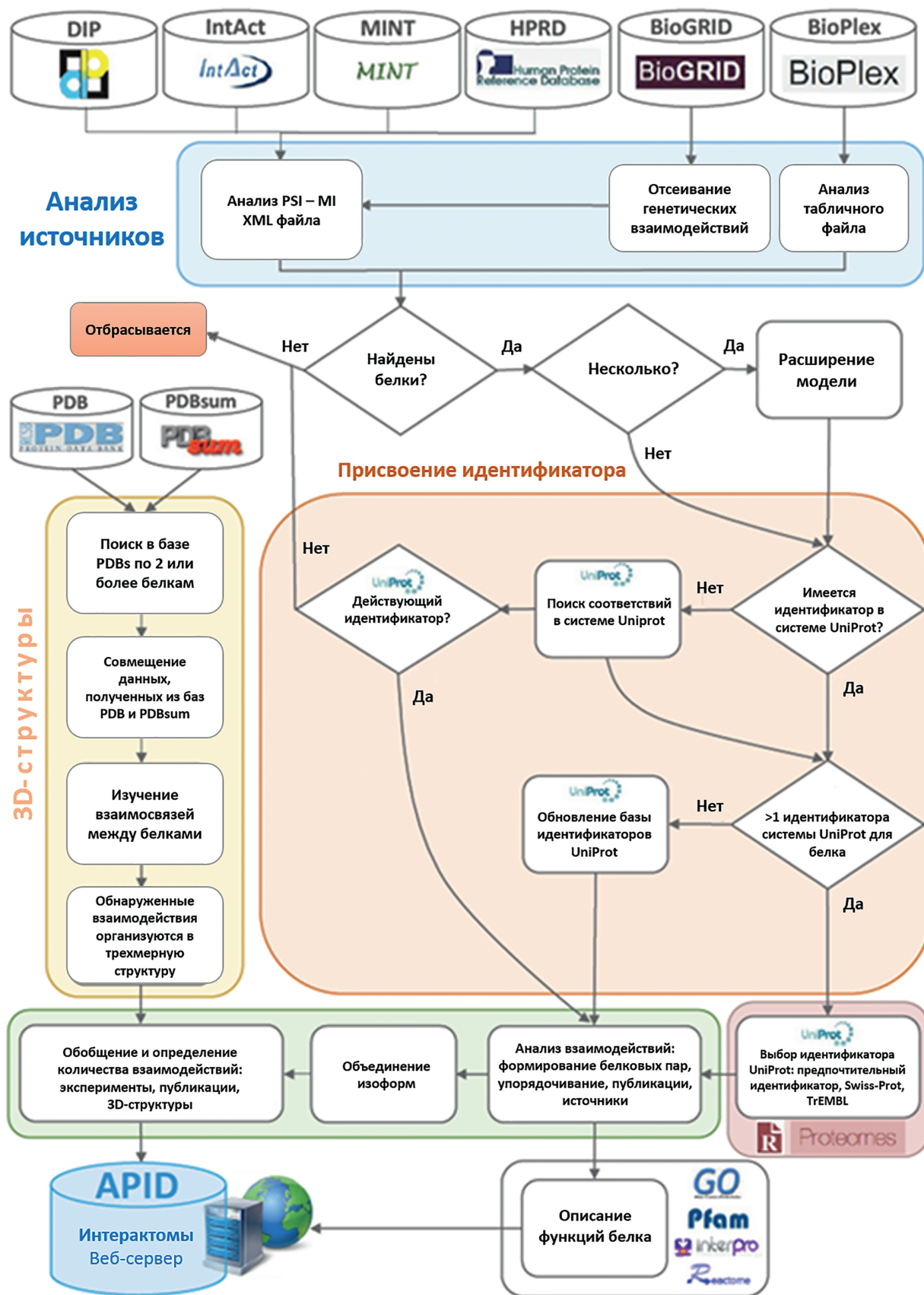


Рис. 5. Общая схема методологии и процедур, используемых для построения APID (Agile Protein Interactomes Data Server): основной рабочий процесс с конвейерами и шагами, применяемые для интеграции данных о межбелковых взаимодействиях (<http://ciclblade.dep.usal.es:8080/APID/init.action#subtab2>)

**Интерактомы на разных уровнях качества и покрытия протеомов
71 785 белков в протеоме**

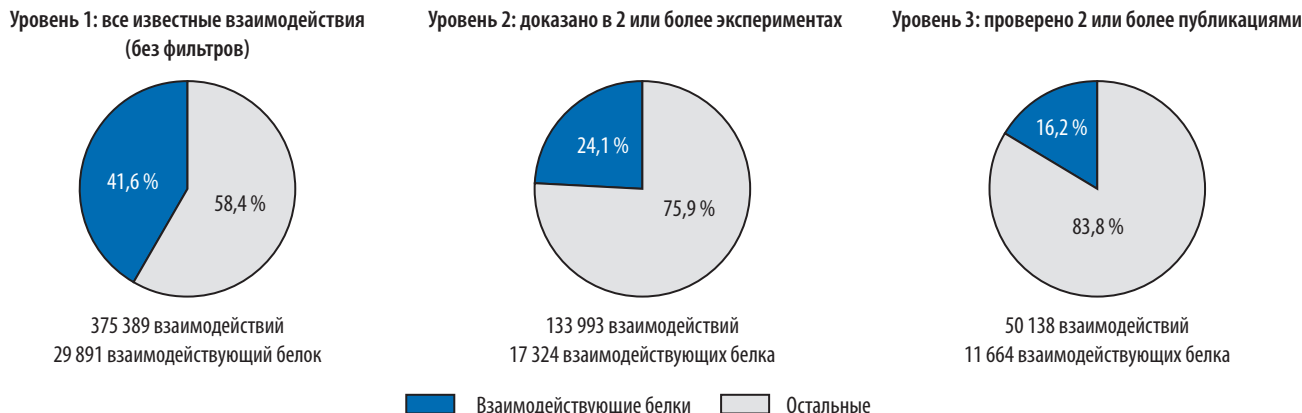


РИС. 6. Данные интерактома *Homo sapiens* на сервере APID. Левая диаграмма показывает, что в настоящий момент по крайней мере 29 701 белок (42,1 % всех белков) участвует как минимум в 349 144 взаимодействиях, которые отражены на интерактомных картах. В противоположность этому количество взаимодействий, подтвержденных в двух независимых экспериментах или в двух отдельных публикациях, существенно меньше (диаграммы в *середине* и *справа*)

тификации и обработке всех белков и их отображению в виде эталонных протеомов каждого вида. Было использовано рекурсивное отображение белка или генного идентификатора в AC/ID UniProtKB в качестве ключевого способа интеграции и унификации данных. Это позволяет избежать дублирования и неправильной идентификации белков.

STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) представляет собой биологическую базу данных и веб-ресурс известных и прогнозируемых белок-белковых взаимодействий.

База данных STRING содержит информацию из многочисленных источников, включая экспериментальные данные, данные литературы и расчеты прогнозируемых взаимодействий. Она находится в свободном доступе и регулярно обновляется. Этот ресурс также предоставляет возможность анализировать функциональные обогащения в представляемых пользователем списках белков с использованием ряда функциональных систем классификации, таких как GO, Pfam и KEGG. Последняя версия (10.0) содержит информацию приблизительно о 9,6 млн белков из более чем 2000 организмов. STRING была разработана консорциумом академических учреждений, включая CPR, EMBL, KU, SIB, TUD и UZH. По состоянию на 2017 г. база данных STRING содержит белок-белковые интерактивные сети с 1380 млн взаимодействий при участии 9,6 млн белков из 2031 организмов (<https://string-db.org/>).

Интерактомы могут различаться в тканях и по стадиям развития. Большинство интерактомов неполные. Экспериментальные методы изучения интерактомов страдают ложноположительными взаимодействиями, что приводит к тому, что до 30 % всех взаимодействий могут являться артефактами [18].

Интерактомы эффективны в выявлении модулей взаимосвязанных генов, которые играют важную роль в заболеваниях. Как и в социальных сетях, расстояние между узлами в интерактоме предполагает высокую или низкую вероятность совместной активности. Таким образом, интерактомы являются полезными инструментами для характеристики биологических процессов на глобальном уровне и для генерирования новых гипотез, которые могут быть дополнительно проверены экспериментально и применены при создании лекарств.

Исследователи из Гарвардского университета в США недавно опубликовали экспериментально подтвержденный интерактом более чем 25 % всего генома человека [19]. Этот интерактом, который получил название BioPlex 2.0, был составлен методом co-IP + MS. Он содержит 56 553 взаимодействия для 10 961 белка. Марковская модель для кластеризации взаимодействующих белков выявила более 1300 белковых хаба, представляющих разнообразные биологические клеточные активности. Например, белками, необходимыми для клеточной адаптации, обогащены 53 хаба, представляющих центральные клеточные функции. Более того, были идентифицированы 442 хаба, относящиеся к 2053 болезням, в результате чего многочисленные белки и гены, связанные с этими болезнями, были объединены в интерактом (<http://bioplex.hms.harvard.edu/index.php>).

ИНТЕРАКТОМ И БОЛЕЗНИ

Генетические болезни и вызывающие их гены изучаются с помощью сцепленного наследования и полногеномных ассоциативных исследований.

В результате этих исследований выявляются локусы хромосом, в которых могут находиться несколько сотен генов кандидатов, потенциально связанных с болезнями. Появление технологии секвенирования всего генома позволило сузить число генов-кандидатов. Однако наличие большого количества нуклеотидных полиморфизмов в генах с неизвестными функциями может затруднить идентификацию драйверного гена, вызывающего болезнь. Использование сетей белковых взаимодействий может помочь для определения приоритетов генов в интервалах сцепления. Для этого сначала выполняется виртуальное выделение (virtual pull-down) белка. При этом для каждого белка-кандидата в данном интервале генома запрашивается база данных на наличие потенциальных взаимодействующих белков-партнеров. Таким образом, для каждого белка-кандидата создается виртуальный комплекс. Далее фенотипы, известные для компонентов комплексов, и фенотип, связанный с интервалом сцепления, сравниваются, и выбираются приоритетные гены путем оценки сходства фенотипов (рис. 7).

Альтернативный метод комбинирования полногеномных ассоциативных исследований (ПАГ)

с сетями белковых взаимодействий был предложен Калифано и соавт. [20], которые утверждают, что просчитывание сигналов ассоциации с учетом биохимических путей улучшает отношение сигнал—шум и увеличивает предсказательную способность ПАГ.

Исследования интерактома при сахарном диабете. Недавнее исследование на группе больных сахарным диабетом 2-го типа (type 2 diabetes, T2D) позволило определить 420 характерных интерактомных мотивов (interactome signatures) сахарного диабета [21]. В указанном исследовании результаты транскрипционных изменений в клетках крови у больных сахарным диабетом были интерпретированы в комбинации полногеномных ассоциативных исследований (Genome-Wide Association Studies, GWAS). Результаты исследования позволили создать интерактом T2D, который объясняет сложные регуляторные механизмы патогенеза этой болезни. Такой интерактом обладает большим потенциалом предсказывать функции и фармакологическую ценность отдельных компонентов системы. Интерактом T2D позволил выявить новый локус риска заболевания сахарным

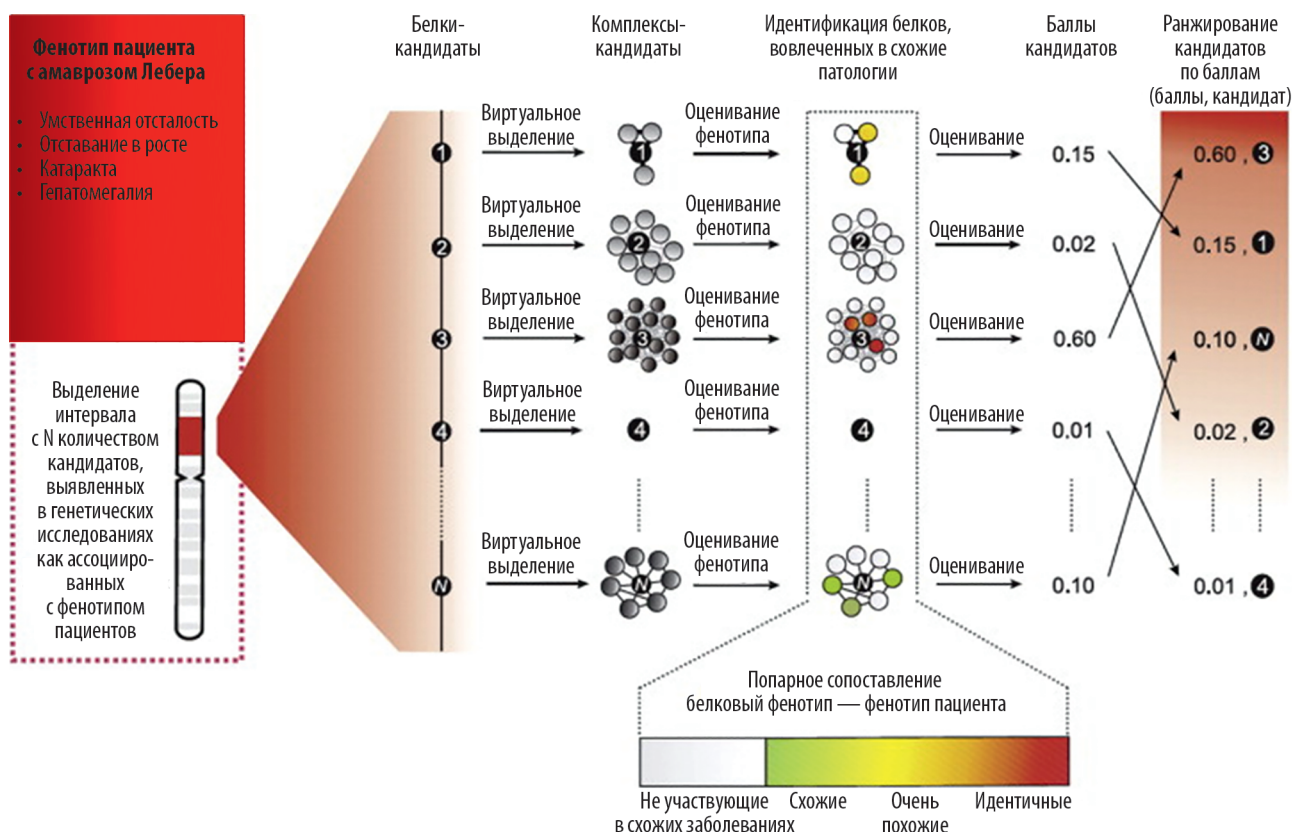


РИС. 7. Использование белковых комплексов для определения приоритетов генов в интервалах сцепления. В данном примере белки, которые участвуют в различных наследственных болезнях, сравнивают с врожденным амврозом Лебера (ВАЛ, наследственное заболевание сетчатки). Белки-кандидаты окрашены в соответствии с их клиническим сходством с фенотипом ВАЛ. Последний шаг заключается в подсчете и ранжировании кандидатов байесовским классификатором (класс алгоритмов классификации). Каждый кандидат оценивается на основе фенотипов, связанных с белками в комплексе-кандидате, и все кандидаты в интервале ранжируются на основе этой оценки. (Цит. по [9], с изм.)

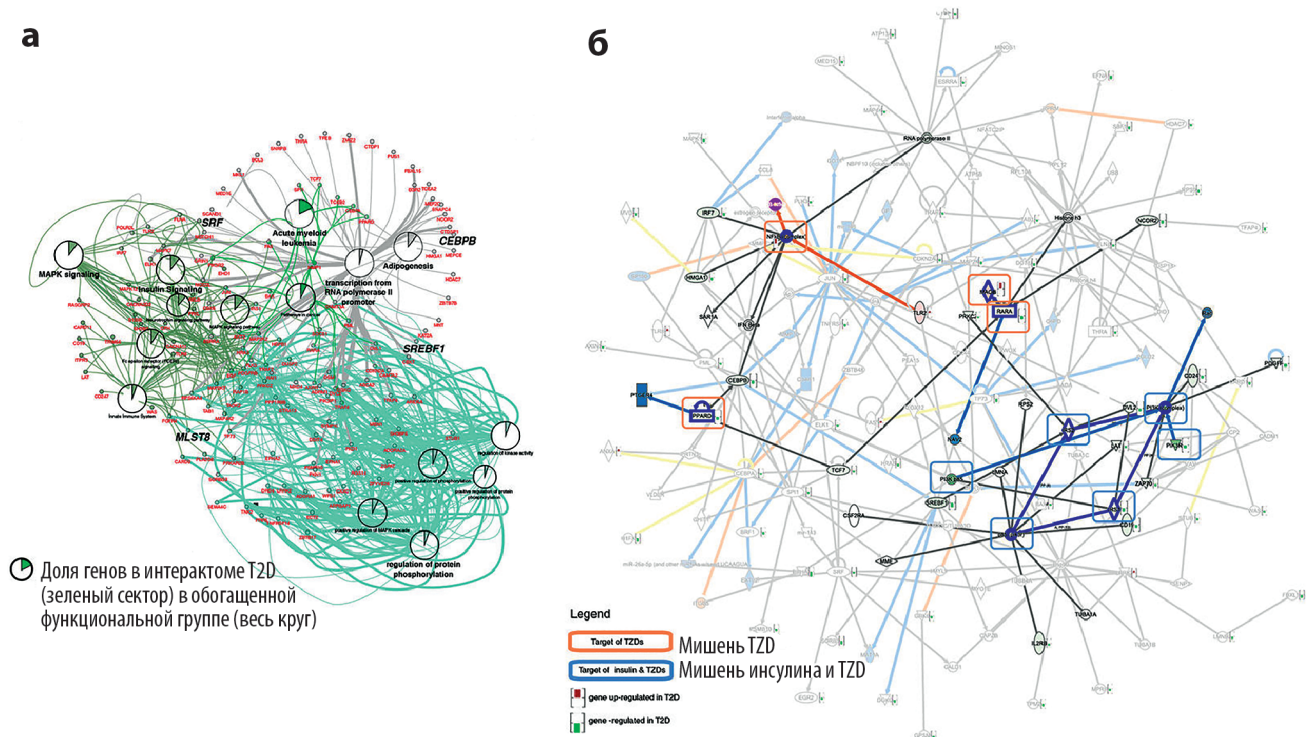


РИС. 8. (а) Функциональный анализ интерактома T2D выявил обогащение его генами, относящимися к сигнальным системам инсулина и MAPK, а также генами, контролирующими транскрипцию, адипогенез и фосфорилирование белков. (б) Варианты интерактомных мотивов T2D, которые подвергаются воздействию антидиабетических лекарств тиазолидинона (TZD) и инсулина. (Цит. по [21], с изм.)

диабетом в геноме (*AChE*) (рис. 8, а), а также указать на известные гены, которые могут быть мишенями фармакологического воздействия (*PPARD*, *MAOB*) с целью лечения этой болезни (рис. 8, б).

Исследования интерактома при шизофрении. Шизофрения имеет ярко выраженную генетическую природу. Рабочая группа при Психиатрическом геномном консорциуме (Psychiatric Genomics Consortium, или PGC) проанализировала полногеномные ассоциативные исследования 36 989 пациентов с шизофренией и 113 075 здоровых индивидуумов и идентифицировала 108 геномных локусов, связанных с риском заболевания шизофренией [22]. Хотя в ходе этого исследования роль генетики в заболевании была четко подтверждена, функциональная роль отдельных рассчитанных вариантов риска оказалась недостаточно понятной.

Исследования с использованием интерактомов показали хороший прогностический потенциал в анализе функциональной ассоциации генов. Например, функция ассоциированного с шизофренией гена *DISC1* первоначально была неизвестна, хотя домены его белкового продукта, такие как спираль-поворот-спираль и лейциновая застёжка, были хорошо охарактеризованы [23]. После определения белок-белковых взаимодействий с использованием Y2H-технологии исследователи успешно связали

DISC1 с системой действия вторичного мессенджера cAMP, процессом удлинения аксонов и миграцией нейронов [24] (рис. 9). Обычно такие исследования проводятся в уже известных сетях взаимодействия белков или, как в случае *DISC1*, когда уже имеется конкретно интересующий ген (или белок). При этом межбелковые взаимодействия определяются такими методами, как Y2H-технология дрожжей или аффинная очистка с последующей масс-спектрометрией.

Другие исследователи, изучая интерактом при шизофрении, обнаружили 57 новых связей для 17 белков. Причем ранее не имелось данных о совместном функционировании этих молекул [25]. Из 57 найденных белок-белковых взаимодействий 19 являются новыми фармакологическими мишенями для 130 используемых лекарственных препаратов.

Интерактом рака. Рак возникает из нормальных тканей вследствие воздействия различных внешних и внутренних факторов. В результате во время многоэтапного процесса канцерогенеза из-за накопления множественных генетических и эпигенетических изменений клетки приобретают злокачественные фенотипы. Гетерогенность соматических мутаций в раковых клетках и неоднородность эпигенетического ландшафта в клетках различных видов опухолей препятствуют построению унифицированных сетей межбелковых взаимодействий.

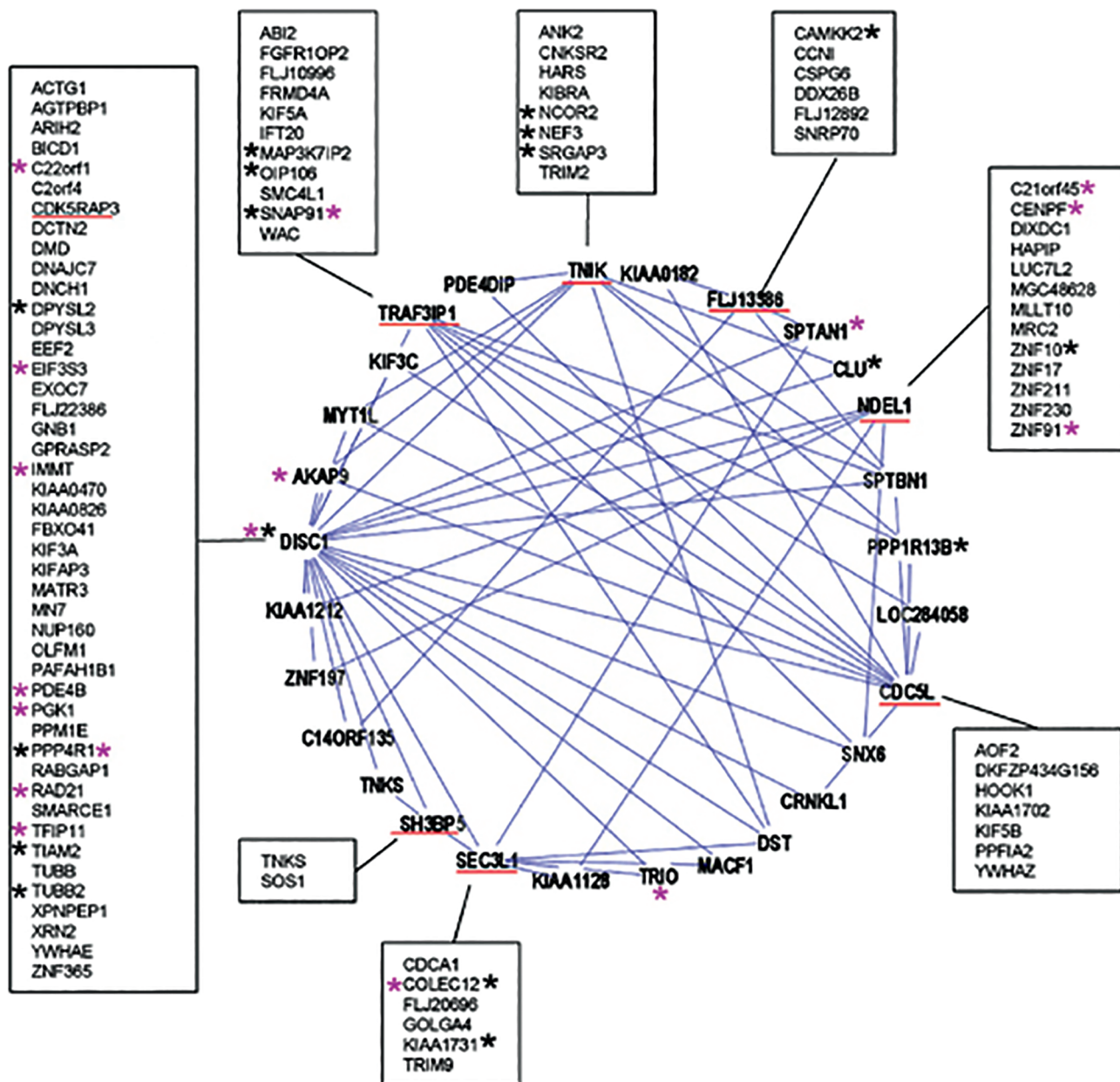


РИС. 9. *DISC1* в сети межбелковых взаимодействий (интерактома). Красным подчеркнуты белки, которые взаимодействуют с *DISC1*. Синие ребра указывают на отношения между молекулами, которые взаимодействуют по крайней мере с еще двумя белками в сети. В рамках: черные звездочки обозначают белки, расположенные в локусах риска развития шизофрении; красные звездочки указывают на белки, которые обнаруживаются в регионах, где хромосомные аномалии связаны с шизофренией. (Цит. по [24], с изм.)

Недавно был предложен оригинальный метод просчитывания функционального интерактома рака, интегрирующего в себе данные из баз полногеномных ассоциаций, соматических мутаций и фармакологических мишеней [26]. Группа исследователей из США и Тайваня подвергли анализу общедоступные данные из таких крупных источников, как база GWAS, данные соматических мутаций при раке и данные о лекарственных препаратах, применяя специально адаптированный сетевой подход. Такой подход позволил выявить, что три выбранных

класса, связанных с заболеванием узлов, демонстрируют неслучайные топологические характеристики в контексте функционального взаимодействия. В частности, было показано, что мишени лекарственных средств, как правило, лежат перед соматическими и герминальными мутациями. Кроме того, был опробован новый подход к измерению иерархии между фармакологическими мишенями, соматическими мутантами и генами восприимчивости к болезням, используя информацию о направленности и длине ребра. В целом подход дает новое

представление о внутренних взаимосвязях между различными классами узлов и расширяет наше понимание природы рака.

Группа канадских ученых обнаружила наличие двух классов хабов в модулярной архитектуре человеческого интерактома [27]. Один класс демонстрирует низкую корреляцию коэкспрессии между генами-партнерами, так называемые межмодульные хабы. Второй класс, называемый внутримодульными хабами, имеет более высокую корреляцию в коэкспрессии между генами-компонентами данного хаба. Например, большинство компонентов протеасомы 26S имеют высокую степень коэкспрессии, чтобы обеспечивать эффективное функционирование системы расщепления белков. Было обнаружено, что межмодульные хабы содержат больше сигнальных белков, чем внутримодульные хабы. Эти результаты свидетельствуют о том, что внутримодульные и межмодульные хабы демонстрируют различные структурные характеристики, соответствующие их ролям в организации коммуникации и функционировании динамических белковых сетей. Сигнальные сети межмодульных хабов обеспечивают функциональную взаимосвязь внутримодульных хабов. Корреляционный анализ по Пирсону, проведенный на 79 биопсиях опухолей молочной железы, показал, что мутации в генах межмодульных хабов чаще ассоциировались с фенотипом рака, чем мутации в генах внутримодульных хабов [28]. Было продемонстрировано, что данный метод может иметь прогностическую ценность для больных раком молочной железы. Были изучены хабы с достоверными

различиями в коэффициенте корреляции Пирсона (PCC) между пациентками, которые были вылечены от рака молочной железы (далее называемого «благоприятным исходом», или good outcome), и пациентами, которые скончались в результате болезни («неблагоприятный исход», или poor outcome). Такой подход выявил наличие 256 хабов, которые отображали изменение PCC как функцию результата заболевания. Одним из таких центров был ген *BRCA1*, который часто мутирован в подгруппе раковых опухолей молочной железы (Breast Cancer Susceptibility Protein 1). Экспрессия *BRCA1* высоко коррелировала с экспрессией ее интерактомных партнеров у выживших пациентов, но не коррелировала с их экспрессией в опухолях пациентов с летальным исходом. Из партнеров по коэкспрессии с *BRCA1* высокая корреляция с благоприятным исходом наблюдалась для *MRE11* и *BRCA2*. Эти белки являются членами супер-комплекса (BRCA1-associated genome surveillance complex, или BASC) [29]. Данный комплекс включает в себя опухолевые супрессоры и белки системы репарации ДНК. Результаты анализа показали, что дезорганизация BASC за счет потери координированной совместной экспрессии компонентов прямо коррелирует с неблагоприятным исходом при раке молочной железы (рис. 10).

ЗАКЛЮЧЕНИЕ

Исходя из предположения, что интерактомы лежат в основе отношений генотипа и фенотипа, свойства

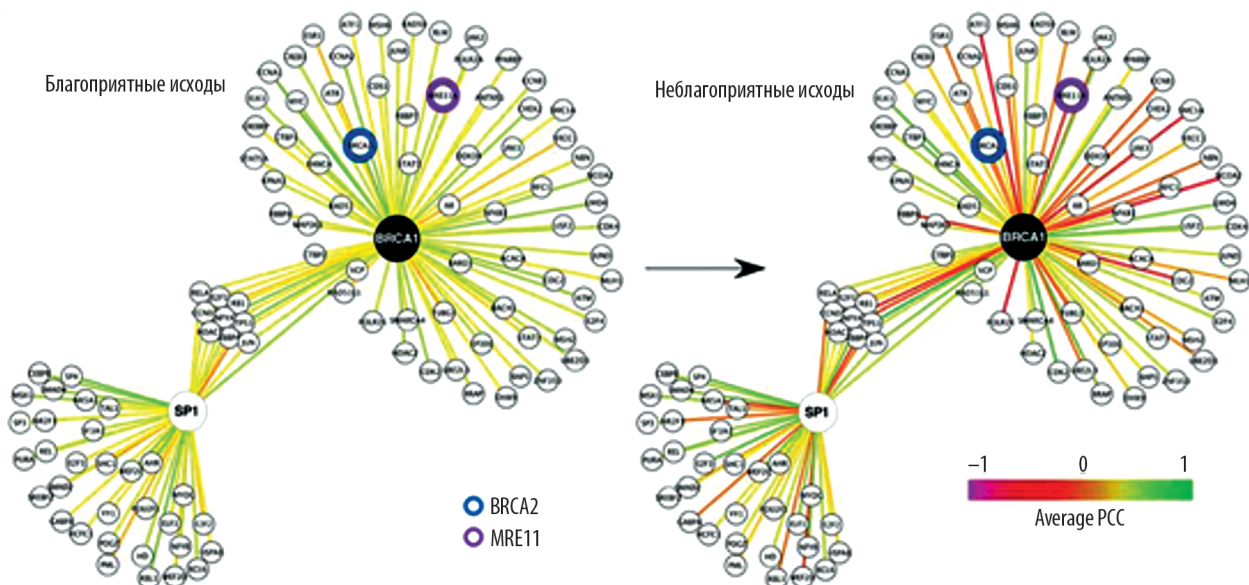


РИС. 10. Различия в динамических свойствах интерактома при опухолях молочной железы. Экспрессии генов *BRCA1* и его взаимодействующих партнеров *BRCA2* и *MRE11* с высокой степенью коррелируют у выживших пациентов (good outcome). Зеленые ребра указывают на высокую корреляцию в экспрессии между парами. Эта корреляция нарушена (красные ребра) у пациентов, которые умерли в результате рака молочной железы (poor outcome). PCC — коэффициент корреляции Пирсона. (Цит. по [27], с изм.)

сети интерактома должны иметь прогностический характер для еще не идентифицированных генетических заболеваний. На примере, приведенном выше, было показано, что интеграция сети коэкспрессии хорошо известных генов, связанных с раком молочной железы, совместно с генетическими и физическими взаимодействиями формирует модель интерактомной сети рака молочной железы, из которой могут быть предсказаны потенциальные гены восприимчивости к онкозаболеванию и гены-модификаторы рака.

Учитывая высокий уровень сложности, неудивительно, что интерактомные сети еще не полностью просчитаны. Предполагается, что принципы картирования интерактомов будут усовершенствованы за счет усложнения и повышения качества наборов данных. В скором времени станет доступен полный каталог вариантов последовательностей среди приблизительно 7 миллиардов человеческих геномов, живущих на Земле людей. Эта информация на многие десятилетия обеспечит продолжение революции в биологии в целом и в трансляционной медицине в частности. Наблюдаемые до сих пор феномены, в том числе и те, которые еще не раскрыты, возможно, приобретут более глобальный смысл в огромном контексте информации, охватываемой геномом человека. Биоинформатика и, в частности, интерактомика будут иметь решающее значение для развития трансляционной, прогностической и персонализированной медицины.

Финансирование. Работа не имеет финансовой поддержки.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА/REFERENCES

1. *Cohrs R.J., Martin T., Ghahramani P., et al.* Translational Medicine definition by the European Society for Translational Medicine. *New Horizons in Translational Medicine*. 2015; 2(3): 86–88.
2. *Batman A.M., Miles M.F.* Translating Alcohol Research: Opportunities and Challenges. *Alcohol Res.* 2015; 37(1): 7–14.
3. *Global Translational Regenerative Medicine Market Prospects 2017–2027.* URL: <https://www.visiongain.com/Report/1843/Global-Translational-Regenerative-Medicine-Market-Prospects-2017-2027>.
4. *Beadle G.W., Tatum E.L.* Genetic Control of Biochemical Reactions in Neurospora. *Proc. Natl. Acad. Sci USA*. 1941; 27(11): 499–506.
5. *Stumpf M.P., Thorne T., de Silva E. et al.* Estimating the size of the human interactome. *Proc. Natl. Acad. Sci USA*. 2008; 105(19): 6959–6964.
6. *Goh K.I., Cusick M.E., Valle D., et al.* The human disease network. *Proc. Natl. Acad. Sci USA*. 2007; 104(21): 8685–8690.
7. *Vidal M., Cusick M.E., Barabasi A.L.* Interactome networks and human disease. *Cell*. 2011; 144(6): 986–998.
8. *Snider J., Kotlyar M., Saraon P., et al.* Fundamentals of protein interaction network mapping. *Mol. Syst. Biol.* 2015; 11(12): 848.
9. *Lage K.* Protein-protein interactions and genetic diseases: The interactome. *Biochim. Biophys. Acta*. 2014. 1842(10): 1971–1980.
10. *Xenarios I., Salwinski L., Duan X.J., et al.* DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids. Res.* 2002; 30(1): 303–305.
11. *Kerrien S., Aranda B., Breuza L., et al.* The IntAct molecular interaction database in 2012. *Nucleic Acids. Res.* 2012; 40: D841–D846.
12. *Ceol A., Chatr Aryamontri A., Licata L., et al.* MINT, the molecular interaction database: 2009 update. *Nucleic Acids Res.* 2010; 38: D532–D539.
13. *Chautard E., Fatoux-Ardore M., Ballut L., et al.* MatrixDB, the extracellular matrix interaction database. *Nucleic Acids Res.* 2011; 39: D235–D240.
14. *Goll J., Rajagopala S.V., Shiao S.C., et al.* MPIDB: the microbial protein interaction database. *Bioinformatics*. 2008; 24(15): 1743–1744.
15. *Breuer K., Foroushani A.K., Laird M.R., et al.* InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Res.* 2013; 41: D1228–D1233.
16. *Niu Y., Otasek D., Jurisica I.* Evaluation of linguistic features useful in extraction of interactions from PubMed; application to annotating known, high-throughput and predicted interactions in I2D. *Bioinformatics*. 2010; 26(1): 111–119.
17. *Orchard S., Kerrien S., Abbani S., et al.* Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. *Nat. Methods*. 2012; 9(4): 345–350.
18. *Kiemer L., Cesareni G.* Comparative interactomics: comparing apples and pears? *Trends Biotechnol.* 2007; 25(10): 448–454.
19. *Huttlin E.L., Bruckner R.J., Paulo J.A., et al.* Architecture of the human interactome defines protein communities and disease networks. *Nature*. 2017; 545(7655): 505–509.
20. *Califano A., Butte A.J., Friend S., et al.* Leveraging models of cell regulation and GWAS data in integrative network-based association studies. *Nat. Genet.* 2012; 44(8): 841–847.
21. *Li J.-W., Lee H.-M., Wang Y., et al.* Interactome-transcriptome analysis discovers signatures complementary to GWAS Loci of Type 2 Diabetes. *Scientific Reports*. 2016; 6: 35228.
22. *Schizophrenia Working Group of the Psychiatric Genomics C.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014; 511(7510): 421–427.
23. *Millar J.K., Christie S., Anderson S., et al.* Genomic structure and localisation within a linkage hotspot of Disrupted In Schizophrenia 1, a gene disrupted by a translocation segregating with schizophrenia. *Mol. Psychiatry*. 2001; 6(2): 173–178.
24. *Camargo L.M., Collura V., Rain J.C., et al.* Disrupted in Schizophrenia 1 Interactome: evidence for the close connectivity of risk genes and a potential synaptic basis for schizophrenia. *Mol. Psychiatry*. 2007; 12(1): 74–86.
25. *Ganapathiraju M.K., Thahir M., Handen A., et al.* Schizophrenia interactome with 504 novel protein–protein interactions. *Npj Schizophrenia*. 2016; 2: 16012.
26. *Ung M.H., Liu C.C., Cheng C.* Integrative analysis of cancer genes in a functional interactome. *Sci Rep*. 2016; 6: 29228.
27. *Taylor I.W., Linding R., Warde-Farley D., et al.* Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat. Biotechnol.* 2009; 27(2): 199–204.
28. *Brown K.R., Jurisica I.* Online predicted human interaction database. *Bioinformatics*. 2005; 21(9): 2076–2082.
29. *Wang Y., Cortez D., Yazdi P., et al.* BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev.* 2000; 14(8): 927–939.