

Карты болезней — фармацевтика *in silico*



ОТ БАЗ ДАННЫХ — К НОВЫМ ЗНАНИЯМ

Созданная в 80-х годах XX в. база генетической информации GenBank сегодня содержит данные приблизительно о 200 млн нуклеотидных последовательностях 100 000 живых организмов. Для обнаружения специфических последовательностей ДНК, характерных для определенных заболеваний, сравнивают геномы больших групп больных и здоровых людей. В 2012 г. в результате сопоставления геномов нескольких тысяч человек было проанализировано 63 524 915 020 пар генетических маркеров и обнаружено множество новых ассоциаций, которые могут служить биомаркерами биполярного расстройства, ишемической болезни сердца, артериальной гипертензии, воспаления кишечника, ревматоидного

Колоссальные объемы накопившейся информации об активности генов, взаимодействиях молекул и реакциях метаболизма в человеческом организме все чаще структурируют в виде баз данных открытого доступа. Этому способствуют рост вычислительных мощностей компьютерной техники и разработка новых алгоритмов информационных технологий. На смену методологиям *in vivo* и *in vitro* пришли компьютерные модели *in silico*

артрита и сахарного диабета. Проект Национальных институтов здравоохранения США «От баз данных — к новым знаниям» (Big Data to Knowledge) с 2013 г. поддерживает биомедицинские исследования и разработку инновационных технологий для крупномасштабного анализа данных.

IN SILICO

Стремительное развитие информационных технологий позволило создать совершенно новую среду для исследований. Если раньше компьютер рассматривали только как средство для обработки экспериментальных данных, то теперь виртуальное пространство полноценно используют в качестве принципиально новой среды для проведения экспериментов. Выражение *in silico* (кремний — основа полупроводников) дополнило классические определения характера экспериментальной среды — *in vivo* (на живых организмах) и *in vitro* (в пробирке). С помощью метода *in silico* и используя соответствующие базы данных, в биоинформатике решают задачи по моделированию поведения отдельных молекул, биохимических процессов и функционированию физиологических систем. Молекулярное моделирование включает набор вычислительных методов, позволяющих изучать системы разной сложности, а молекулярный докинг подразумевает изучение взаимодействия типа лиганд — рецептор или фермент — субстрат. Моделирование на основании гомологии позволяет определять пространственную структуру белковых молекул с учетом имеющихся структурных шаблонов

белков с гомологичными аминокислотными последовательностями.

СЕТИ МОЛЕКУЛЯРНЫХ ВЗАИМОДЕЙСТВИЙ

Определение характера взаимодействий между молекулами в живом организме настолько важно для фармацевтики и многих естественных наук, что было выделено отдельное исследовательское направление, получившее название «интерактомика». Соответственно, интерактоном называют полный набор (схему) межмолекулярных взаимодействий в отдельной клетке. Таким образом, интерактомика позволяет взглянуть на организм или биосистему в целом. В ней собраны большие наборы полногеномных и протеомных данных и определены корреляции между молекулами. Из данной информации формулируют новые гипотезы об ответных реакциях этих молекул, которые могут быть проверены в новых экспериментах. Чаще всего термин «интерактом» относится к сетям белок-белковых взаимодействий или к их подмножествам. Белок-ДНК интерактом называют гено-регуляторной сетью, поскольку он представлен сетью, образованной транскрипционными факторами, хроматин-регуляторными белками и их генами-мишенями. Метаболические сети тоже рассматривают как сети молекулярных взаимодействий, в которых химические компоненты клетки превращаются друг в друга с помощью ферментов, связанных со специфическими субстратами. Все типы интерактонов взаимосвязаны и перекры-

ваются. Например, белковые интерактомы содержат много ферментов, которые в свою очередь формируют биохимические сети. Подобным же образом ген-регуляторные сети пересекаются с сетями белковых взаимодействий и с сигнальными сетями. В результате изучения этих сетей формируются функциональные карты происходящих в клетке процессов, идентифицируют мишени лекарств и выявляют новые функции генов.

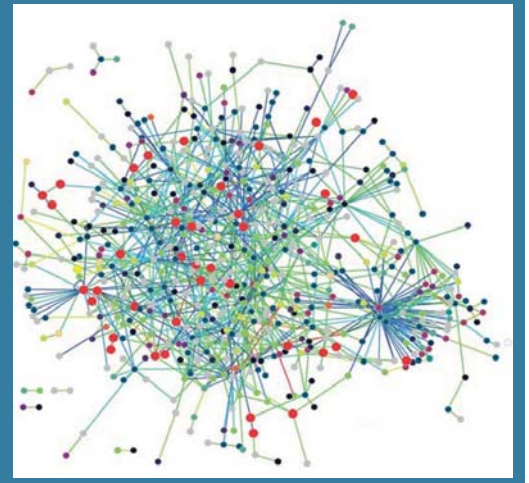
ПОЛНЫЙ ИНТЕРАКТОМ

Возможно, когда-нибудь мы все узнаем о молекулярных взаимодействиях в клетках человека. Существует множество баз интерактомных данных, среди которых следует особо отметить проект интерактома человека и базу данных человеческих протеиновых связей, предсказанных *in silico*. В 2010 г. наиболее полный из существующих интерактомов, основанный на 5,4 млн попарных сравнений генов, описывал взаимодействие 75% всех генов и включал 170 тыс. взаимодействий. Построив полный интерактом (схему молекулярных взаимодействий в клетке), можно будет определить, на что повлияет поломка каждого из белков или генов. На данный момент, по оценкам ученых, нам известно около 20% интерактома человека, и при этом он исследован неравномерно. Но даже неполная карта дает возможность извлечь полезную информацию о заболеваниях. Это самый эффективный метод для раскрытия роли уже известных и прогнозирования особенностей функционирования не описанных ранее генов. Модель позволяет рассматривать генетические взаимодействия в разных масштабах. Со временем благодаря уточнению карты интерактома ее можно будет использовать для прогнозирования эффектов тех или иных мутаций и раскрытия новых терапевтических возможностей.

КЛАСТЕРЫ ДЕФЕКТНЫХ БЕЛКОВ

Чаще всего заболевание вызывается взаимодействием нескольких типов молекул и обусловлено «поломкой» нескольких видов генов. При этом замечено, что мутантные белки, провоцирующие заболевание, часто взаимодействуют между собой. В таком случае на карте интерактома выделяют кластеры (модули) поврежденных белков, характерных для каждой болезни. На неполной карте интерактома пока различимы только крупные кластеры заболеваний, в патогенезе которых участвует много белков (не менее 25). На сегодня известно несколько сотен патологических состояний, связанных с изменениями большого количества белков. На актуальной карте интерактома ученым удалось обнаружить

Определение характера взаимодействий между молекулами в живом организме настолько важно для фармацевтики и многих естественных наук, что было выделено отдельное исследовательское направление, получившее название «интерактомика»



кластеры для 226 заболеваний. Однако некоторые из связанных с болезнями белков все же оказались вне кластеров. К примеру, с рассеянным склерозом ассоциировано 69 белков, но лишь 11 из них были связаны между собой и образовывали кластер. Тем не менее связанных белков во всех случаях было больше, чем можно было ожидать при случайном распределении. Примечательно, что в кластеры объединяются в среднем около 20% связанных с болезнью белков — на данный момент известна именно такая доля всех взаимодействий. Расположение патогенных белков на карте интерактома помогает выявлять новые связи между уже известными данными. Когда карта молекулярных взаимодействий станет более подробной, ученые рассчитывают увидеть на ней гораздо больше связанных кластеров.

СВЯЗИ МЕЖДУ ЗАБОЛЕВАНИЯМИ

Кластеры белков, участвующих в патогенезе различных заболеваний, на карте интерактома могут располагаться на разном удалении друг от друга и даже перекрываться. Их расположение раскрывает новую информацию о связях между заболеваниями, а расстояние между ними отражает сходство симптомов болезней, а также вероятность того, что оба заболевания разовьются у одного человека. Если в развитии нескольких болезней задействованы одни и те же белки, то кластеры могут перекрываться. К примеру, в патогенезе бронхиальной астмы и целиакии, связанной с непереносимостью глютена и нарушением всасывания в тонкой кишке, неожиданным образом задействован определенный набор общих белков. Оказалось, что для этих непохожих диагнозов также характерен общий однонуклеотидный полиморфизм — единичная замена нуклеотида в ДНК. По-видимому, между этими патологиями

есть много общего. Для таких заболеваний с перекрывающимися кластерами, как рассеянный склероз и ревматоидный артрит, обнаружены общие патобиологические механизмы и симптомы. Но ученые выявили и другие случаи перекрывания кластеров, когда у сходных по проявлениям состояний не было общих патогенных белков. Эти случаи особенно интересны, поскольку обнаружить связь между заболеваниями просто по спискам задействованных молекул с помощью других методов невозможно. Кроме того, по мере возрастания степени перекрывания кластеров увеличивалась и вероятность возникновения этих заболеваний одновременно. Одним из неожиданных результатов стало перекрывание кластеров лимфомы и инфаркта миокарда, в патогенезе которых не было выявлено одинаковых молекул. Однако, как оказалось, эти состояния достаточно часто встречаются вместе. Кроме того, ранее было установлено, что при внутрисосудистой В-крупноклеточной лимфоме опухолевые клетки могут нарушать проходимость сосудов, в том числе и в миокарде, вызывая острую ишемию. Пересечения кластеров на карте интерактома могут выявить общие закономерности даже между не связанными на первый взгляд болезнями. Такие связи в перспективе помогут установить механизмы развития менее изученных заболеваний на основе данных о более изученных.

Именно в сфере биоинформационных технологий совершаются новые фундаментальные открытия. Интеграция биологии, медицины и фармацевтики наряду с тесным взаимодействием с информационными методологиями обещает удивительные результаты.

Татьяна Кривомаз, д-р техн. наук, канд. биол. наук, профессор