

Многомерная биология XXI века и клиническая лабораторная диагностика

Вельков В.В.
 ЗАО «ДИАКОН»

Одна из главных задач XXI века – как остановить пандемическое распространение болезней цивилизации: атеросклероза, ишемической болезни сердца, сахарного диабета, ожирения, метаболического синдрома, онкологических заболеваний. Что должна сделать лабораторная диагностика для решения этой задачи? Во-первых, уметь своевременно определять генетическую предрасположенность к возникновению наиболее серьезных патологий. Как становится все более очевидным, многие такие патологии могут вызваться генетическими мутациями. Во-вторых, с высокой достоверностью определять количественный показатель риска возникновения патологий на доклинической стадии. Это позволит своевременно разработать и практически реализовать программы по предупреждению развития заболеваний. И, в-третьих, за счет динамического измерения новых биомаркеров осуществлять мониторинг эффективности проводимых мероприятий. Эти задачи уже успешно решаются. И происходит это благодаря революционным достижениям биологии XXI века – ее назвали многомерной биологией (high dimensional biology). В нее входят:

Геномика – идентификация *всех* генов человека и нарушений в них, приводящих либо к наследственным заболеваниям, либо к предрасположенности к ним.

Транскриптомика – идентификация *всех* матричных РНК, кодирующих белки, определение количества каждой индивидуальной мРНК, определение закономерностей экспрессии всех генов, кодирующих белки.

РНОмика – идентификация *всех* не кодирующих РНК, определение количества каждой индивидуальной нкРНК – определение закономерностей экспрессии всех нкРНК.

Метаболомика – идентификация и количественное определение *всех* метаболитов, синтезируемых (или находящихся) в данных клетках, тканях, органах и в биологических жидкостях.

Биоинформатика. Использование вычислительной техники, математики и информационной теории для анализа и моделирования молекулярно-биологических систем, в особенности систем, состоящих из генов, РНК, белков и метаболитов и др. Создание баз данных.

Именно эти направления, для краткости называемые OMICS (*genomics, transcriptomics* и т.д.), считаются основой медицины XXI века.

Клиническая геномика и клиническая транскриптомика.

Транскрипционные профили патологий

Реализация проекта «Геном человека» позволяет обнаруживать мутации в генах, приводящие к наследственным заболеваниям или к повышению вероятности возникновения многих патологий, таких, например, как онкологические, сердечно-сосудистые (атеросклероз), сахарный диабет, метаболический синдром, шизофрения и др. В практику лабораторной диагностики уже успешно внедрены методы идентификации мутаций, наиболее часто приводящих, например, к различным раковым заболеваниям. Эти методы основаны на применении полимеразной цепной реакции (ПЦР) и маркерных генов, содержащих нарушения, приводящие к конкретным патологиям. Развитие геномики патологий позволяет, однако, не только проводить их молекулярно-генетическую диагностику, но и, как следующий этап, определять интенсивность синтеза РНК и белков, имеющих отношение к возникновению и развитию заболеваний. Это осуществляется с помощью определения транскрипционных профилей, характеризующих экспрессию всех генов, активных в данном образце. Технологии, применяемые для решения этих задач основаны на так называемых «ДНК микрочипах» (DNA microarray). Генный чип представляет собой твердую подложку, на которую в определенном порядке нанесены в виде точек индивидуальные гены (их ДНК). Чтобы определить, транскрибируется ли данный ген,

на чип помещают (с определенными координатами) лишь его часть – олигонуклеотид. Этот олигонуклеотид соответствует экспрессируемой части гена (экзону). Затем из образца (например, опухоль) выделяется вся (суммарная) РНК. На основе всех молекул РНК данного образца получают их ДНК-копии – кДНК (обратная транскрипция), которые флуоресцентно метят и потом проводят гибридизацию с иммобилизованными на микрочипе олигонуклеотидами. Если в данных условиях какие-то точки с конкретными генами не гибридизуются – это значит, что данный ген не транскрибируется. Если же данная точка микрочипа «светится», значит олигонуклеотиды на этой площадке прогибридизовались с флуоресцентно меченой кДНК, ген транскрибируется.

Чтобы определить, является ли полученный результат ошибкой или нет, проводится сравнение двух объектов. Для этого берут образец А (патология), из него получают суммарную РНК и после обратной транскрипции всех ее молекул флуоресцентно метят определенным цветом (красным) все молекулы кДНК. То же проводят и с образцом В (норма), но метят молекулы кДНК другим цветом (зеленым). Затем проводят гибридизацию ДНК-микрочипа со смесью этих двух препаратов кДНК (конкурентная гибридизация – преимущественно образуют гибриды те молекулы, которых больше). Если сигнал в данной точке на чипе будет красным, значит в клетках А (патология) транскрипция данного гена сильнее, чем в клетках В (норма). Если сигнал зеленый, то транскрипция сильнее в клетках В (норма). Если красного и зеленого поровну, то получится желтый цвет. Таким образом, можно сравнивать уровень транскрипции данного гена в разных тканях и органах, в биологических жидкостях при норме и патологии, до терапии и в ее процессе, до хирургической операции и после. Довольно часто термины «геномика», «транскриптомика» и «протеомика» употребляются в одном и том же значении – для обозначения анализа экспрессии всех генов данного образца – как на уровне синтеза мРНК, так и на уровне синтеза белков [1–3].

Транскриптом – набор **всех** РНК, находящихся в данном образце. Анализ транскриптома, определение качественного и количественного профиля всех синтезированных РНК, отражает синтез кодируемых ими белков, а так же синтез рибосомальных, транспортных и других РНК. Сравнение транскриптомов нормальных и патологических образцов, позволяет идентифицировать новые маркеры, проследить изменение их

уровней во времени, судить о динамике патологии, об эффективности проводимого лечения и прогнозировать его результат [4–6]. Предполагается, что каждая болезнь, характеризуется своим, так сказать, «штрих-кодом» – уникальным паттерном уровней транскрипции набора генов, характерного именно для данной болезни. Анализ транскриптом осуществляют с помощью компьютерных методов распознавания образов.

Клиническая «РНомика»

Пожалуй, самой громкой сенсацией биологии конца XX стало открытие принципиально нового класса РНК. Практически во всех эукариотных организмах неожиданно было обнаружено огромное количество различных РНК, которые **не кодируют белков** и не являются ни рибосомальными, ни транспортными. Они играют, в основном, регуляторную роль – влияют на экспрессию генов (чаще всего на уровне трансляции). Это, так называемые, микроРНК длиной 19–22 нуклеотида, к ним относятся и т.н. интерференционные РНК, которые выключают синтез определенных белков путем разрушения их мРНК. МикроРНК регулируют экспрессию генов после их транскрипции. Это может происходить за счет: 1) репрессии трансляции мРНК, 2) расщепления мРНК, 3) ускорения распада мРНК. В каждой микроРНК есть участок, комплементарный особому участку в той мРНК, которая при каких-то обстоятельствах подлежит инактивации. Таким образом, большинство мРНК имеют «черные метки», указывающие на возможность собственной деградации, а микро РНК, имеющие комплементарные участки, в нужный момент узнают «черные метки» и нацеливают на мРНК, приговоренные к ликвидации, ферменты и белки, для этого предназначенные. С помощью РНомики микроРНК, содержащиеся в образцах, идентифицируются, определяется их концентрация. Оказалось что, изменения концентрации различных микроРНК характерно для различных типов патологий. И в клинической РНомике наступила «золотая лихорадка» по поиску «золотых» маркеров и предикторов заболеваний [7–9].

Клиническая протеомика

Клиническая протеомика – это идентификация и количественное определение всех индивидуальных белков, которые содержатся в биологическом материале (сыворотка крови, спинномозговая жидкость, моча, биопсия) и мониторинг изменения их концентраций. Протеом – совокупность **всех** белков, содержащихся в данном образце. Полный анализ протеома клеток, тканей, органов

и биологических жидкостей проводится с помощью двумерного электрофореза с высоким разрешением и с последующей идентификацией индивидуальных белков за счет масс-спектрометрии. Это позволяет проанализировать до 10 000 индивидуальных белков в одном образце и зафиксировать изменения их концентраций. Типичная последовательность операций при протеомике такова: 1) отбор образца (клетки, ткань, биологическая жидкость), 2) приготовление образца (лизис клеток, экстракция белков), 3) изоэлектрофокусировка, электрофорез в 1-ом направлении, 4) электрофорез во 2-ом направлении (полиакриламидный гель, додецилсульфат натрия), 5) проявление белковых пятен на геле, 6) анализ двумерной электрофореграммы (количество пятен, их расположение), 7) выделение участков геля, содержащих индивидуальные белковые пятна, 8) расщепление индивидуальных белков трипсином прямо в геле, 9) масс-спектрометрический анализ: а) масс-фингерпринтинг пептидов, б) определение аминокислотных последовательностей фрагментов индивидуальных белков, 10) идентификация каждого белка и измерение его концентрации, документирование, обработка результатов.

Значительному прогрессу в области протеомики способствовали успехи масс-спектрометрического анализа пептидов. Масс-спектрометрия включает в себя четыре основных компонента. Во-первых, в ионном источнике масс-спектрометра из образца получают ионизированные пептиды или белки. Во-вторых, разделение ионов пептидов и белков происходит в анализаторе масс на основе их величины отношения массы к заряду (m/z). В-третьих, детектор ионов (время пролетного масс-спектрометра), регистрирует отдельные ионы, с указанием значения m/z иона, количества ионов, и времени пролета ионов от источника ионов до детектора ионов. И, наконец, интерпретация полученных данных с помощью биоинформатики, анализ баз данных, в итоге, получение дифференциального профиля белков. С помощью этой технологии уже открыты новые белковые маркеры и получены впечатляющие результаты в области **кардиоваскулярной** протеомики и **онкопротеомики** [10–13].

Кардиоваскулярная геномика, транскриптомика и протеомика

Повышать вероятность возникновения и развития атеросклероза могут мутации во многих генах. Такие мутации идентифицированы, например, в гене ALOX5AP, который кодирует белок активатор 5-липоксигеназы (5 lipoxygenase 5-LOX). Этот белок активирует синтез

лейкотриена – липида, медиатора воспалительного процесса в стенках сосудов. Также повышают вероятность возникновения атеросклероза мутации в гене MEF2A, который кодирует фактор регуляции транскрипции в миоцитах; в гене Apo E, который кодирует аполипопротеин E, входящий в состав липопротеинов высокой плотности (ЛПВП); в гене LTA, кодирующего синтез альфа-лимфотоксина [12–13]. Уже созданы базы данных по сотням белков протеома миокарда, уровни которых изменяются при хронических и острых сердечно-сосудистых заболеваниях. Обнаружено, что при этих заболеваниях возрастают концентрации т.н., белков теплового шока (heat-shock proteins – HSP), белков митохондрий, а так же белков, вовлеченных в генерирование энергии. Диагностически важные белки кардиопротеома классифицируются так: 1) белки, связанные с получением энергии и метаболизмом, 2) белки, индуцируемые стрессом, 3) белки, обеспечивающие сократительные функции и формирование цитоскелета [14].

Весьма перспективным оказался мониторинг динамики протеомов тканей, взятых при помощи биопсии до и после хирургического вмешательства. Обнаружено более 100 кардиоспецифических белков, уровни которых значительно изменяются при дилатационной кардиомиопатии. Практически при всех формах сердечной недостаточности, ее начальная стадия – это компенсаторная адаптация сердечной ткани к патологическим изменениям – в частности, гипертрофия левого желудочка. Довольно часто в таких случаях, клинические, электрокардиографические и гемодинамические показатели не достаточно адекватно отражают переход от патологии к норме. Однако возвращение к норме при вентрикулярном ремоделировании может быть эффективно диагностировано при мониторинге кардиопротеома [15]. Согласно предварительным данным, наиболее обещающими маркерами для кардиопротеомики могут быть тропонины, изоформы альфа-1-фибриногена, изоформы аполипопротеина А-1, С-реактивный белок и др. В передовых клиниках хирурги уже сейчас читают кардиопротеомы так же уверенно, как электрокардиограммы и ангиограммы [16].

Транскриптомика и протеомика плазмы крови

Сколько **всего** индивидуальных белков в плазме крови? Какие новооткрытые белки плазмы могут иметь диагностическое значение? Когда и как изменяются их уровни? Проект «Протеом Плазмы Крови» (ППК) был начат

в 2002 году. В реализации ППК участвуют 35 лабораторий в 13 странах. На первой стадии проекта была создана аннотированная база данных **3020** белков плазмы крови человека [www.bioinformatics.med.umich.edu/hupo/ppp; www.ebi.ac.uk/pride]. Она содержит информацию о: 1) идентификации каждого белка, 2) его концентрации и стабильности, 3) об особенностях его масс-спектрального анализа и др. [17, 18]. На второй стадии проекта было идентифицировано уже **9504** белка (на основе масс-спектрометрической идентификации одного или двух пептидов каждого белка) и **3020** белков (при идентификации двух и более пептидов, что более точно). **889** белков плазмы идентифицированы с достоверностью в 95%. Следующий этап – построение протеомов плазмы, характерных для различных патологий и формулировка международно согласованных диагностических показателей [19, 20].

Протеомика гемостаза

Протеомика тромбоцитов – быстро развивающееся направление. Уже обнаружены ранее неизвестные белки тромбоцитов, быстро идет расшифровка механизмов, приводящих к нарушениям коагуляции. В ранних исследованиях протеомов тромбоцитов было обнаружено, что при активации они секретируют 82 белка. Затем при использовании тромбоцитов, стимулированных тромбином, было установлено, что **секретом** (секретируемый протеом) тромбоцитов – это более 300 белков и только 37% из них были известны ранее. Примерно 35% белков, секретируемых тромбоцитами, синтезируются и другими клетками. Только 28% белков, секретируемых тромбоцитами, обнаруживаются в местах атеросклеротических повреждений. Эти работы привели к идентификации новых мишеней для новых лекарственных препаратов, таких как секретогранин III, циклофилин А, калуменин, а также пролили свет на возможные механизмы адгезии тромбоцитов, способствующие развитию атеросклероза [21].

Онкогеномика, онкотранскриптомика, онкопротеомика

Впечатляющие достижения современной биологии в области ранней диагностики и лечения онкологических заболеваний. Обнаружены маркерные гены, активирующиеся на ранних стадиях онкогенеза и соответствующие им маркерные белки. Найдены маркеры, позволяющие проводить молекулярную классификацию опухолей, обнаружены предикторы метастазирования,

предикторы ответа на терапию, маркеры, позволяющие прогнозировать развитие различных онкозаболеваний. Основные задачи онкопротеомики таковы: 1) построение протеомов и анализ их динамики при возникновении и развитии различных опухолей; 2) идентификация путей передачи клеточных сигналов, приводящих к онкогенезу; 3) идентификация маркеров для диагностики онкозаболеваний и для мониторинга ответа опухоли и организма на хирургическое вмешательство и на разные типы терапии, 4) определение иммунного ответа на онкогенез [22–25].

Основная проблема во внедрении онкопротеомики в клиническую практику – сложность обучения онкологов чтению карт онкотранскриптомов и онкопротеомов.

Как уже отмечалось выше, одно из самых сенсационных открытий в молекулярной биологии, сделанных в конце XX века, обнаружение микроРНК. На данный момент у человека идентифицировано около 400 генов, кодирующих разные микроРНК и этот список будет возрастать. Интерес к ним крайне высокий. Как показывают результаты последних лет, изменения в синтезе микроРНК сильно связаны с возникновением, прогрессированием и метастазированием злокачественных опухолей. Часть микроРНК при этом синтезируется в больших количествах, а синтез других падает. Некоторые исследователи даже полагают, что именно нарушение регуляции синтеза микроРНК, которые являются регуляторами синтеза белков – если не первопричина онкогенеза, то, по крайней мере, одна из главных причин [26, 27].

Более 50% генов, кодирующих известные микроРНК человека, расположены в областях хромосом, связанных с онкогенезом. Некоторые микроРНК могут функционировать как онкогены – индуцируют онкогенез. К этому приводит повышение их синтеза. Другие микроРНК проявляют себя как супрессоры опухолей – подавляют неконтролируемую пролиферацию. Например, избыточный синтез микроРНК *mir-17-92*, которая проявляет себя как онкоген, подавляет активность гена, который, в свою очередь, должен обеспечивать синтез белка – супрессора опухолей или белка, стимулирующего апоптоз («запрограммированную смерть») раковых клеток. В свою очередь сниженный синтез некоторых микроРНК, например, микроРНК *let-7*, проявляется как действие опухолевого супрессора, способного ингибировать онкогенез за счет инактивации белков, вызывающих деление клеток [28]. Отсюда и название микроРНК, связанных с онкогенезом – **oncomirs (oncogenic micro RNA)** [29].

С помощью онкоРНОмики идентифицирован, в частности, комплекс микроРНК, который позволяет однозначно дифференцировать рак поджелудочной железы и доброкачественные опухоли этого органа. В этот комплекс входят около 100 различных микроРНК. Их содержание в опухолях поджелудочной железы в 30–50 раз выше нормы. Открытие этих микроРНК не только повысит возможности ранней диагностики рака поджелудочной железы, но и, возможно, ляжет в основу создания препаратов, ингибирующих их активность и, тем самым, подавляющих развитие опухолей поджелудочной железы [30].

Повышенные уровни микроРНК miR-103 и miR-107, сопровождающиеся исчезновением микроРНК miR-155, позволяют проводить дифференциальную диагностику опухолей эндокринных желез и ацинозных опухолей. Повышенный синтез микроРНК miR-204 связан с инсулиномами и коррелирует с повышенным уровнем инсулина, регистрируемым иммуногистохимическими методами [31].

Использование знаний о микроРНК перспективно не только для диагностики онкологических заболеваний. Предполагается, что введение в раковые клетки синтетических или природных РНК, предназначенных для избирательного подавления патологических повышенных уровней онкомикроРНК – весьма перспективный метод молекулярной терапии злокачественных заболеваний. Работы в этом направлении ведутся весьма интенсивно. Ожидается, что в 2010 году мировой рынок терапевтических препаратов, созданных на основе микроРНК, составит 3,5 млрд долл, а в 2015 – 10,5 млрд долл. [32].

Эндокринная транскриптомика и протеомика

Различные эндокринные клетки синтезируют разные спектры как кодирующих РНК, так и микроРНК. Весьма интенсивно ведутся работы по транскриптомике и протеомике **всех** эндокринных органов, как в норме, так и при патологиях. Результаты регулярно поступают в базу данных «Омнибус Экспрессии Генов» (Gene Expression Omnibus – GEO, <http://www.ncbi.nlm.nih.gov/geo/>) [33]. Цель эндокринной транскриптомики и протеомики: установить цепь молекулярных событий, от индукции синтеза гормона до реализации его действия в норме и при патологиях, обнаружить мутации, влияющие на эти процессы и разработать методы идентификации таких мутаций, применимые в лабораторной практике. Уже проводится идентификация **всех** РНК, изменение

синтеза которых свидетельствует о злокачественных заболеваниях эндокринной системы [34–35]. В 2002 г было идентифицировано около 400 генов, кодирующих рецепторы, связанные с различными G-белками и взаимодействующие с гормонами [36]. Весьма перспективен транскриптомный и протеомный мониторинг результатов гормональной терапии многих онкологических заболеваний. При разработке методов мониторинга гормональной терапии эстроген-зависимого рака груди был установлен комплекс из 138 генов, активность которых изменяется в ответ на действие эстрогена [37].

Не менее перспективны результаты исследований, полученные в области перинатальной транскриптомики и протеомики, аллергопротеомики, геномики, транскриптомики и протеомики стресса.

Клиническая метаболомика

Метаболомика – построение глобального профиля концентрации **всех** метаболитов в данном образце. Основное направление работ – выявление метаболических изменений, характерных для инициации патологий и ее динамики, для закономерностей ответа метаболизма на терапию. Основные патологии, находящиеся в фокусе метаболомики – метаболический синдром, сахарный диабет, сердечно-сосудистые заболевания, патология печени. Метаболомика основана на применении спектроскопии протонного ядерного магнитного резонанса в сочетании с компьютерным анализом распознавания образов. Распознавание образов (паттернов) мультиплетной структуры спектров ЯМР – это новый инструмент изучения структуры и свойств органических соединений. Для лабораторной диагностики определяющее значение имеют спектры протонного ядерного магнитного резонанса сыворотки, спинномозговой жидкости и мочи [38–41]. Метаболомика уже показала высокую эффективность при обнаружении врожденных и наследственных нарушений метаболизма, нарушений метаболизма, вызванных эндогенными и экзогенными факторами, при трансплантациях (до и после), при изучении токсичности лекарственных препаратов (**токсикогеномика**), реакций организма на лекарственные препараты (**фармакогеномика**), при определении индивидуальных особенностей реакции организма на различные пищевые продукты (**нутриномика**) [42].

Клиническая липидомика – важнейшее направление метаболомики. Нарушения липидного обмена связаны с такими заболеваниями, как атеросклероз, сахарный

диабет, ожирение, болезнь Альцгеймера. **Липидом** – липидный профиль грубого липидного экстракта из образца – это масс спектр, характеризующий липидный состав и концентрации **всех** индивидуальных липидов образца. Подход основан на комбинации жидкостной хроматографии и масс-спектрометрического анализа. Прогресс в липидомике достигнут благодаря разработке новых масс-спектрометрических подходов, в частности, методов «мягкой ионизации», таких как ионизация электрораспылением и матрично-активированная лазерная десорбция/ионизация. В частности, успехи применения метода масс-спектрометрии с ионизацией электрораспылением способствовали развитию «струйной» липидомики (shotgun lipidomics) и практическому применению разделения компонентов внутри источника ионов в качестве стратегии применения двумерной масс-спектрометрии для изучения состава липидов биологических образцов [43].

Липидомика подразделяется на: 1) липидомiku клеточной архитектуры и мембран (architecture/membrane lipidomics) и, 2) липидомiku медиаторов (mediator lipidomics). С помощью липидомики строятся метаболические сети, в которых участвуют (практически) **все** липиды и медиаторы, являющиеся производными липидов. С помощью этого подхода уже установлено детальное строение мембран многих типов клеток, установлены механизмы активации провоспалительных цитокинов, происходящие за счет медиаторов, связанных с липидами [44, 45].

Нарушения метаболизма липидов связано с серьезными неврологическими патологиями, включающими биполярные расстройства и шизофрению, а также с нейродегенеративными заболеваниями, такими как болезнь Альцгеймера, болезнь Паркинсона, болезнь Неймана–Пика. Особенно большое значение имеет нейролипидомика спинномозговой жидкости [46].

Многомерная биология: перспективы для медицины и лабораторной диагностики

Научное значение многомерной биологии для медицины переоценить невозможно. Этот подход уже привел к выявлению новых механизмов возникновения и развития многих патологий. В ближайшем будущем следует ожидать появления единой многомерной медицинской науки, раскрывающей детальные молекулярные механизмы конкретных патологий в цепи событий, включающих: гены → РНК → белки → метаболиты →

биохимические и физиологические процессы → психиатрические и психические особенности → ментальные и интеллектуальные характеристики.

Что касается ценности «... *омики*» для лабораторной диагностики – это открытие в внедрение в клиническую практику новых маркеров патологий. Диагностическая ценность маркеров обычно зависит от трех показателей: чувствительности, специфичности и предсказательной способности. Как правило, между чувствительностью и специфичностью существует обратная зависимость. К сожалению, маркеры с идеальной специфичностью и чувствительностью очень редки. Но с другой стороны, причина почти всех патологий никогда не бывает единственной, чаще всего это неблагоприятное сочетание многих отрицательных факторов. Возможным решением этой проблемы может стать разработка стандартных комплексов маркеров, которые будут давать достоверную предсказательную и/или диагностическую информацию. Поскольку биология высоких измерений позволяет установить причинно-следственные связи в цепи: ген → РНК (кодирующая или микроРНК) → белок → метаболит, такие комплексы маркеров могут состоять из специфических олигонуклеотидов, белков и метаболитов [47].

Следующий этап – разработка для относительно широкого применения узкоспециализированных наборов для транскриптомики, протеомики и метаболомики, например, кардиопатологий, ренальных патологий, различных злокачественных заболеваний, разных биологических жидкостей и т.д., (реагенты, оборудование, инструменты, компьютеры, базы данных, программное обеспечение и др.). Это позволит создать компьютерную модель больного. Появится возможность лечить одновременно больного и его компьютерную модель, построенную на основе геномики, транскриптомики, протеомики и метаболомики пациента и, по мере лечения, на основе мониторинга их динамики, корректировать и то другое.

И тогда, возможно, медицина XXI века станет предиктивной, превентивной и персонализированной: предсказывающей, предотвращающей и ориентированной не на борьбу с отдельными болезнями, а на четкое и научно индивидуализированное лечение конкретного больного.

* Список литературы находится в редакции