

УДК 577.21

## ПРОТЕОГЕНОМИКА ЕДИНИЧНЫХ КЛЕТОК – БЛИЖАЙШАЯ ПЕРСПЕКТИВА

### Обзор

© 2020 С.А. Мошковский<sup>1,2\*</sup>, А.А. Лобас<sup>3</sup>, М.В. Горшков<sup>3</sup>

<sup>1</sup> Российский национальный исследовательский медицинский университет им. Н.И. Пирогова,  
117997 Москва, Россия; электронная почта: smosh@mail.ru

<sup>2</sup> Научно-исследовательский институт биомедицинской химии им. В.Н. Ореховича, 119121 Москва, Россия

<sup>3</sup> Институт энергетических проблем химической физики им. В.Л. Тальрозе,  
Федеральный исследовательский центр химической физики им. Н.Н. Семенова РАН, 119334 Москва, Россия

Поступила в редакцию 15.10.2019

После доработки 11.11.2019

Принята к публикации 11.11.2019

Технические достижения в области геномных технологий последних лет привели к взрывному росту исследований живых систем на уровне единичных клеток в масштабах целых транскриптомов. В обзоре представлено как вслед за транскриптомикой свой путь в анализе единичных клеток начинает протеомика. Уже появились первые работы по использованию хроматомасс-спектрометрического анализа полных протеомов на отдельных клетках. Разделение клеток в них осуществляют по аналогии с транскриптомным анализом, например, методом клеточного сортирования, а масс-спектрометрический анализ проводят с помощью модифицированного метода tandemных массовых меток. Объединение результатов транскриптомного и протеомного анализа в рамках протеогеномного подхода к молекулярному профилированию анализируемых клеток улучшит понимание механизмов клеточного взаимодействия как при развитии организмов, так и в различных патологиях.

**КЛЮЧЕВЫЕ СЛОВА:** протеомика, транскриптомика, протеогеномика, анализ единичных клеток, таргетная массовая метка (ТМТ), масс-спектрометрия.

DOI: 10.31857/S0320972520020025

Молекулярное профилирование с использованием омиксных технологий разного вида [1] до недавнего времени имело дело с биологическими образцами, объединяющими свыше нескольких тысяч, а иногда и миллионов отдельных клеток. С учетом того, что участки тканей, состоящие из морфологически сходных клеток, редко встречаются в организмах, сравнение транскриптомов или протеомов таких образцов обладает принципиальным недостатком, который принято называть «средней температурой по больнице». Особенно неблагоприятно этот эффект сказывается на результатах поиска биомаркеров, где неоднородность представленности тех или иных молекул в клетке может приводить к ложно-положительным результатам. Понятно, что более корректными с точки зрения результата могли быть работы по сравнению биологических образцов большого объема в случае анализа биологических жидкостей [2], где молекулы распределены более или менее равномерно, или

однородных клеточных культур, в особенности, синхронизированных [3]. Частично проблему гетерогенности исследуемых образцов решают микродиссекцией: однородные участки ткани выявляют либо визуально, либо с использованием автоматизированного анализа изображений, получаемых под микроскопом. Затем однородные участки разделяют физическими методами и подвергают молекулярному анализу [4].

Таким образом, для молекулярного сравнения тканей многоклеточных организмов омиксными методами, в том числе и в медицинских целях, существует явная потребность в поиске новых подходов для анализа единичных клеток. Первые успехи в этом направлении были достигнуты в области транскриптомов единичных клеток [5]. В настоящем обзоре обсуждаются проблемы появившихся в самое последнее время подходов к протеомному анализу единичных клеток в контексте достижений транскриптомики в этой области, а также возможности интеграции уже на новом уровне результатов этих двух омиксных технологий.

\* Адресат для корреспонденции.

## МОЛЕКУЛЯРНОЕ ПРОФИЛИРОВАНИЕ ЕДИНИЧНЫХ КЛЕТОК – ТЕХНИЧЕСКИЙ ПРОРЫВ В АНАЛИЗЕ ТРАНСКРИПТОМОВ

Анализ транскриптомов единичных клеток с помощью секвенирования нового поколения (NGS) в последние пять лет интенсивно развивается. Предтечей используемых сейчас методов является флуоресцентное секвенирование РНК *in situ* (FISSEQ), в котором фиксированную на подложке ткань использовали для получения кДНК, которую также привязывали к клеточному матриксу [6]. Поскольку секвенирование на платформе Illumina, по сути, основано на флуоресцентной спектроскопии, его вариант удалось адаптировать для секвенирования полученной кДНК непосредственно на срезе ткани. В итоге получали достаточно разрешенные изображения по уровню каждого из зарегистрированных транскриптов, подходящие для сравнения интактных клеток и тканей в различных состояниях [7].

Практически сразу после появления первых работ начался взрывообразный рост разработки методов транскриптомики единичных клеток, и уже через несколько лет существовало не меньше пяти альтернативных технических решений [5], работающих, тем не менее, на примерно одинаковых принципах. Сначала необходимо разделить в пространстве клетки или их ядра. Для этого можно использовать, например, сортировку клеток с помощью флуоресценции (FACS), подобно той, которую давно применяли в клеточных сортерах [8]. В итоге клетки по одной располагаются в планшетах для дальнейшей манипуляций. Сходным образом клеточный сортер от «BD Biosciences» (США) использует разбавление клеточной суспензии, в ходе чего при ее раскапывании в маленькие лунки планшета в каждой из них статистически оказывается по одной клетке [9]. В еще одном широко используемом техническом решении от компании «10X Genomics» (США) клетки в специальном микрофлюидном устройстве по одной связываются с особыми гранулами, а полученные комплексы для дальнейшего анализа упаковываются в отдельные масляные капли [10].

Следующий этап после изоляции клеток – это подготовка РНК к секвенированию, которая, с некоторыми вариациями, напоминает обычную подготовку к NGS, включающую в себя получение кДНК и расщепление ее на считываемые фрагменты. Однако существенным отличием является потребность в индивидуальном мечении каждой клетки, поскольку кДНК от всех выделенных клеток следует вновь объединить для получения высокой производитель-

ности. Поэтому в каждую лунку планшета или на каждую гранулу, используемую в капельной технологии выделения клеток, добавляют индивидуальные, синтезированные комбинаторным путем олигонуклеотидные метки. Впоследствии, после получения огромного набора считываний фрагментов целевой кДНК, по этим меткам уже программным путем распознаются подмножества считываний (ридов), принадлежавшие единичным клеткам [5].

Методы анализа транскриптомов единичных клеток в настоящее время ориентированы, в основном, на анализ кодирующих поли-А-транскриптов. В итоге, в разных вариантах методов, получают полные транскриптомы единичных клеток или т.н. уникальные молекулярные идентификаторы (UMI) генов [11], по которым можно определить их уровень экспрессии в единичных клетках. Предложенные к использованию совсем недавно методы секвенирования транскриптомов единичных клеток конкурируют между собой, а публикации, в основном, сфокусированы на результатах сравнения их эффективности [5, 12].

Вскоре после использования перечисленные выше методы транскриптомики единичных клеток принесли свои плоды в виде ярких, значимых для биологии результатов. В этом обзоре не будет представлен исчерпывающий обзор этих работ, поэтому приведем несколько примеров. Так, по транскриптомам единичных ядер исследователи смогли детализированно проследить за судьбой клеток нервного гребня при их дифференцировке в мышинном эмбрионе [13]. В результате было установлено состояние бинарности, когда клетки находятся в состоянии выбора между двумя траекториями их развития, а также выявлены факторы управления их судьбами. В другой работе было показано, что транскриптомика единичных клеток может определять отношение интронных и экзонных последовательностей в кодирующих белки транскриптах, и был предложен подход для оценки кинетики концентрации индивидуальных РНК (RNA velocity) [14]. Если многие транскрипты содержали интронные последовательности, значит транскрипция только началась, а продукт будет наращивать свою концентрацию. Наоборот, присутствие только зрелых, сплайсированных транскриптов означает конечную фазу экспрессии данного гена. Этот остроумный подход позволил оценивать тенденции экспрессии генов в масштабах единичных клеток по одному «снимку» транскриптома [15]. Особенно важна роль единичных клеток в центральной нервной системе. Выделить нейроны без их разрушения, как представляется, почти невозможно, но тран-

скриптомика единичных клеточных ядер также нашла широкое применение в нейронауках. Транскриптомы единичных ядер позволяют картировать и кластеризовать нейроны коры головного мозга, дополняя другие виды функционального картирования. Сравнение клеток коры человека и мыши позволило выявить существенный консерватизм устройства некоторых ключевых участков, а также выявить видо-специфические черты [16]. Важные эволюционные выводы удалось сделать при анализе единичных клеток головного мозга четырех видов приматов, включая человека [17].

Таким образом, за считанные годы своего существования транскриптомика единичных клеток революционным образом изменила подход к наблюдению за клеточными процессами *in vivo*. Благодаря этому подходу были охарактеризованы многие клеточные типы, которые не нельзя было отличить морфологическими микроскопическими методами [18].

### ПРОТЕОМИКА ЕДИНИЧНЫХ КЛЕТОК

Поскольку амплификация белков, подобная полимеразной цепной реакции для нуклеиновых кислот, недоступна, протеомика отстает от методов анализа нуклеиновых кислот по аналитической чувствительности [19]. В случае протеомики анализ единичных клеток по своей природе предъявляет еще большие требования по этому параметру, в то время как микроскопия и клеточный сортинг, основанные на флуоресценции, уже давно представляли возможность для работы с единичными клетками. Флуоресцентные белки, например, слитые с интересующими продуктами генно-инженерным путем, использовали для исследования накопления этих продуктов в реальном времени, на уровне единичных клеток [20]. К недостаткам этой группы методов относится потребность в генетических манипуляциях, эффекты от которых могут исказить процессы, происходящие в модельных клетках.

Еще более разнообразны методы анализа единичных клеток, основанные на распознавании интересующих клеток антителами и другими специфически связывающими молекулами. Технология клеточного сортирования, основу которой заложили более полувека назад [21], стала самым распространенным методом анализа единичных клеток на белковом уровне [22]. Даже хорошо известный метод Вестерн-блоттинга был адаптирован для анализа единичных клеток, когда их сортируют по микролункам планшета и после лизиса переносят на гель для связывания с антителами [23].

Т.н. масс-цитометрический метод был предложен для анализа отдельных белков на уровне единичных клеток. В этом методе количественные измерения таргетных белков осуществляются с использованием антител, конъюгированных с ионами переходных металлов. Белковые комплексы, содержащие такие антитела, поступают в источник ионизации на основе индуктивно-связанной плазмы (ICP), где происходит ионизация переходных металлов с последующим измерением масс-спектров стабильных изотопов этих металлов с использованием времяпролетного масс-анализатора. С помощью масс-цитометрии была продемонстрирована возможность количественного профилирования иммунного отклика на уровне единичных клеток [24]. В настоящее время этот метод все чаще используется для анализа единичных клеток, например, когда интерес исследователей сфокусирован на ограниченной группе белков. Так, в развивающихся кроветворных клетках с помощью масс-цитометрии проводили количественный анализ ответственных за дифференцировку этих клеток транскрипционных факторов [25].

Следует отметить, что все основанные на антителах методы имеют ряд принципиальных ограничений в случае единичных клеток. Во-первых, при их масштабировании, когда технология становится по-настоящему омиксной, тестирование десятков или даже сотен отдельных антител на предмет перекрестных взаимодействий становится практически невозможным. В основном, антитела проходят отбор на связывание своей мишени, в то время как остается неизвестным, способны ли они связываться структурно сходные мотивы в других белках. Кроме того, антительные методы ограничены проницаемостью клеток и необходимостью диффузии реагента через концентрированное клеточное содержимое [26].

Что же с классической протеомикой, основанной на масс-спектрометрии, где белки и пептиды анализируются напрямую, без применения аффинных реагентов? В режиме панорамной протеомики, направленной на выявление всех белков клетки, с масс-спектрометрическими детекторами высокого разрешения возможна идентификация и полуколичественный анализ белковых продуктов примерно десяти тысяч генов в клеточных линиях, что близко к теоретической оценке всего продуцируемого протеома какого-либо клеточного типа [27]. Однако такой анализ до недавнего времени проводили для протеомов, полученных от десятков и сотен тысяч клеток. Есть ли перспективы у панорамного протеомного анализа единичных

клеток с точки зрения аналитической чувствительности? Действительно, детекторы, используемые в настоящее время в масс-спектрометрии, демонстрируют, по биологическим меркам, высочайшую чувствительность, будучи способными зарегистрировать от нескольких сотен до нескольких тысяч ионов, удерживаемых в ионной ловушке масс-анализатора [28]. При этом, согласно одному из последних мета-анализов, количество молекул белка в дрожжевой клетке составляет ~40 млн [29], а в опухолевой клетке *HeLa* – примерно на два порядка больше [30]. Даже с учетом гетерогенности природных белков, многочисленных потерь на стадиях обработки образца между белком в интактной клетке и регистрируемым ионом пептида от этого белка, которые включают клеточный лизис и трипсинолиз, хроматографию и ионизацию, ситуация с регистрацией протеома единичной клетки методом панорамной протеомики не кажется такой уж безнадежной. Это иллюстрируется исследованиями последних лет, когда было предложено несколько подходов к анализу единичных клеток посредством хроматомасс-спектрометрии, широко принятой в протеомике. Во-первых, проблема чувствительности снимается при анализе очень больших клеток, когда можно использовать обычные методики без существенных изменений. С такими клетками исследователи сталкиваются в эмбриологии. Например, в недавно опубликованном исследовании удалось отследить изменения протеома единичных ооцитов при их созревании *in vitro* [31]. При массе белка на клетку, по оценкам авторов, равной 100 нг, из единичного ооцита удавалось идентифицировать 450 белков. Кроме того, анализировали единичные бластомеры при делении яйца шпорцевой лягушки на стадии крупных бластомеров, в одном удавалось идентифицировать ~1400 белков! [32]. Привлекательной представляется еще не вполне реализованная идея протеомного анализа гигантских нейронов моллюсков, тела которых могут достигать диаметра 400 мкм [33]. При исследовании этого объекта знание о дифференциально продуцирующихся белках может помочь расшифровать важные механизмы работы нервной системы, в частности, раскрыть материальную основу памяти [34].

Гигантские клетки – привлекательная модель для анализа, однако очевидна потребность в исследовании протеома клеток обычных размеров, например, опухолевых или кровяных клеток диаметром 10–30 мкм. Первые шаги в этом направлении уже сделаны. Авторы этих первых работ использовали клеточную сортировку с помощью микрофлюидики, напо-

добие той, которую применяли для транскриптомных исследований единичных клеток. После разделения клетки подвергали обработке, т.н. «тандемными массовыми метками» (TMT, tandem mass tag) [35]. Подход с использованием мультиплексных TMT-меток обеспечивает одновременный количественный анализ нескольких образцов за счет мечения пептидов в каждом из них изотопными метками, которые проявляют себя после фрагментации пептидных ионов в масс-спектрометре. Химические метки сконструированы таким образом, что они имеют одинаковую молекулярную массу, однако после фрагментации пептидов в масс-спектрометре за счет комбинаций встроенных стабильных тяжелых изотопов  $^{13}\text{C}$  и  $^{15}\text{N}$  их фрагменты приобретают различную массу, в зависимости от присоединенной метки. В настоящее время в коммерческих наборах, доступных на рынке, достигается возможность достигать мультиплексности до 16 TMT-меток и проводить количественный анализ их протеомов друг относительно друга (TMT 10-plex, 11-plex, 16-plex, наборы компании «Thermo Fisher Scientific», США).

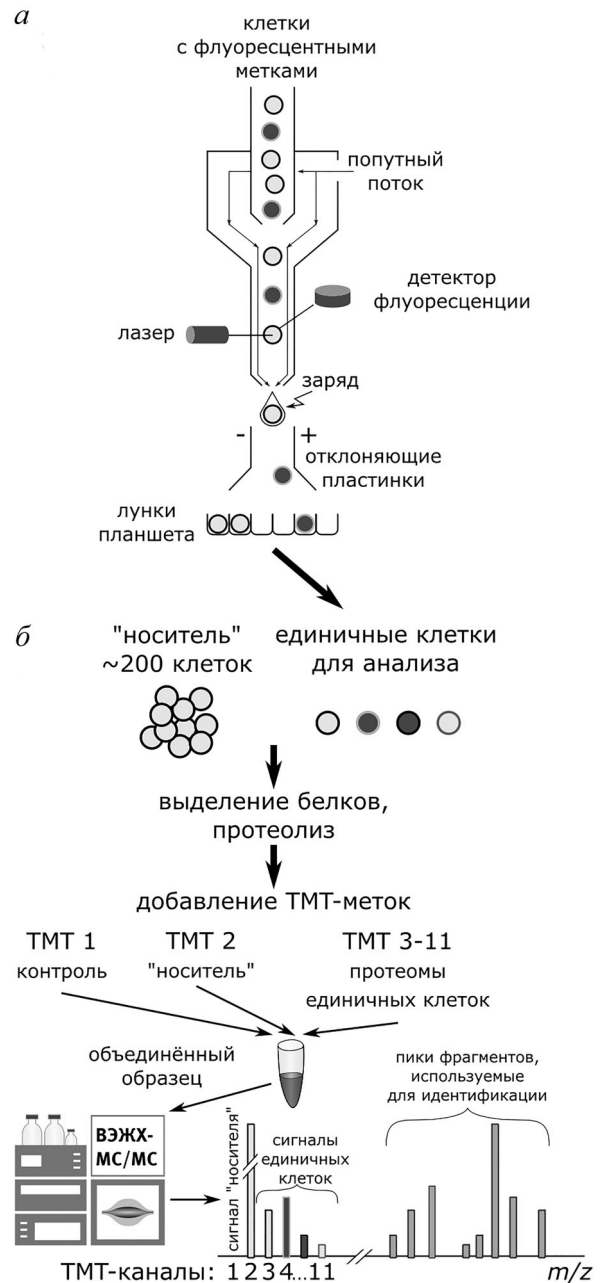
Именно такие наборы TMT-меток использовали Slavov et al. для первой демонстрации панорамной протеомики единичных клеток человека [36]. Метод назвали SCoPE-MS (Single Cell Proteomics by Mass Spectrometry). Клетки разделяли по одной на лазерном сортере (FACS) и помещали в лунки микропланшета (рисунок, а). Одновременно использовали т.н. «носитель» – помещенные в отдельную лунку 200 клеток того же типа. Этот носитель должен обеспечить сигнал для идентификации белков обычным панорамным методом, а единичные клетки должны были сравниваться с сигналом носителя. Действительно, с помощью TMT в первом эксперименте удалось количественно проанализировать более 700 белков в каждой из восьми исследуемых клеток (из 10 TMT-меток одну оставили для контроля, а одной пометили 200 клеток «носителя», рисунок, б). Для работы с белками на пределе чувствительности пришлось изменить настройки масс-спектрометра высокого разрешения на основе ионной ловушки Orbitrap [37] и модернизировать методы обработки первичных масс-спектрометрических данных для идентификации и количественного анализа [38]. За первой, демонстрационной работой по SCoPE-MS последовала следующая, в которой предложенная методика была существенно улучшена с целью повышения производительности [39]. Методом SCoPE2 за 85 ч работы хроматомасс-спектрометра с использованием 11-комплексного набора TMT удалось проанализи-

ровать протеомы 356 единичных моноцитов и макрофагов из иммортализованных клеточных линий. По протеомам, в которых был проведен количественный анализ ~2 тыс. белков, удалось хорошо разделить две клеточные популяции. Безусловно, проблемой использования TMT по сравнению с транскриптомикой является значительно более низкая мультиплексность этого метода и возможность анализа всего 10 клеток одновременно, тогда как при секвенировании РНК все клетки снова смешиваются на чипе секвенатора с надежным разделением результатов программными методами. В целом, результаты, полученные методом SCoPE2, дают повод для осторожного оптимизма.

Другие исследователи также начали работу по анализу единичных клеток с использованием массовых меток и тандемной масс-спектрометрии. Так, в недавней статье описан анализ 72 мышинных клеток с глубиной анализа каждой из них, достигающей до 1600 идентифицированных белков [40]. Процедура заняла около двух суток, что сопоставимо со скоростью анализа моноцитов и макрофагов в ранее упомянутой работе [39]. Клетки разделяли и распределяли по «нанолункам» специального чипа на серийном сортире BD. После этого процедуру пробо-подготовки, включая химическую модификацию белков и гидролиз протеазами, осуществляли с помощью особой роботизированной системы nanoPOTS, где все реакции проходили в каплях нанометрового размера [41].

В процессе подготовки этого материала в открытом доступе появился препринт, описывающий сходное по исполнению исследование единичных клеток линии лейкоза с использованием клеточного сортирования и тандемных массовых меток [42]. Как и в предыдущих работах, хромато-масс-спектры меченых единичных клеток сравнивали с носителем (авторы называли его «boost») из 500 клеток. Для количественного анализа использовали режим получения тандемных масс-спектров (MS2), а также в других запусках прибора получали фрагменты фрагментов трипсиновых пептидов (MS3). Для анализа данных авторы разработали вычислительный конвейер Sceptre, который еще предстоит оценить рецензентам для дальнейшей публикации статьи в журнале.

Очевидно, что методы, позволяющие анализировать протеомы единичных клеток при помощи панорамной масс-спектрометрии, остаются уделом немногих лабораторий, работающих на пределе возможностей технологии, в том числе, с использованием уникальных, сконструированных в лабораторных условиях устройств, в отличие от транскриптомики, где технические



Протеомика единичных клеток с детекцией на основе тандемных массовых меток (TMT), по [36] с изменениями. Сначала клетки с помощью стандартного клеточного сортирования помещают в лунки микропланшета, где их расщепляют протеазами по особому протоколу (а). Затем осуществляют мечение набором TMT, а для сравнения используют образец-«носитель» из 100 или более клеток. Количественный анализ проводят на хромато-масс-спектрометре высокого разрешения (б)

решения для анализа единичных клеток полностью коммерциализированы [5]. Также следует отметить, что упомянутые результаты первых работ по панорамной протеомике единичных клеток [36, 39, 40, 42] пока требуют подтвержде-

ния ортогональными методами, например, с использованием специфичных антител.

### НА ПУТИ К ПРОТЕОГЕНОМИКЕ ЕДИНИЧНЫХ КЛЕТОК

Интеграция данных омических технологий, в случае нуклеиновых кислот и белков обозначаемая как протеогеномика, позволяет более эффективно оценивать молекулярные каскады, вовлеченные в те или иные процессы [43], а также обнаруживать протеоформы и выяснять их функциональность [44]. Здесь и далее, под протеоформами мы понимаем продукты одного и того же гена, различающиеся за счет генного полиморфизма, альтернативного сплайсинга и посттранскрипционных модификаций мРНК, а также протеолитического расщепления и посттрансляционных модификаций белков [45].

В случае протеогеномики единичных клеток существуют два возможных направления развития. Для тех, кто создает уникальные методы, можно задуматься об анализе в рамках одной и той же клетки, и транскриптома, и протеома [46], по аналогии с анализом образцов большого объема, например, биоптатов злокачественных опухолей [47]. Уникальные молекулярные траектории отдельных клеток в таких процессах, как дифференцировка или онтогенез, можно будет рассматривать на двух сочетанных уровнях. Что же касается пользователей стандартизованных протеомных методов, то им, как представляется, также можно воспользоваться разрастающимся банком данных по транскриптомике единичных клеток. Качество секвенирования РНК в некоторых исследованиях уже позволяет производить в транскриптомах отдельных клеток поиск единичных участков полиморфизма (SNP) [48]. Соответственно, в субпо-

пуляциях клеток различных органов и тканей может быть предсказано обогащение несинонимичными SNP, возникшими вследствие мутагенеза или посттранскрипционных модификаций, например, редактирования РНК. Анализ новых протеоформ, предсказанных из данных транскриптомики единичных клеток, может быть осуществлен на группах клеток традиционными таргетными или панорамными протеомными методами. Например, основываясь на данных панорамного протеомного анализа методами хроматомасс-спектрометрии, мы недавно показали, что редактирование РНК ферментами ADAR модифицирует разные белки в глиальных и нейрональных клетках головного мозга мыши [49]. Поиск аналогичных участков редактирования в данных транскриптомики единичных клеток [50] мог бы значительно облегчить понимание функционального значения этого явления.

Таким образом, несмотря на отсутствие амплификации, которая способствует быстрому развитию способов анализа нуклеиновых кислот в единичных клетках, панорамная протеомика также движется к завоеванию своего места в этой области. Несомненным остается то, что оптимальные результаты по анализу молекулярных траекторий единичных клеток в различных процессах могут быть достигнуты при интеграции технологий анализа генной экспрессии и продукции белков.

**Финансирование.** Работа выполнена при финансовой поддержке РФФИ (грант 17-15-01229).

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

**Соблюдение этических норм.** Настоящая работа не содержит описания исследований, выполненных с участием людей или использованием животных в качестве объектов.

### СПИСОК ЛИТЕРАТУРЫ

- Horgan, R.P., and Kenny, L.C. (2011) "Omic" technologies: genomics, transcriptomics, proteomics and metabolomics, *Obstet. Gynaecol.*, **13**, 189–195, doi: 10.1576/toag.13.3.189.27672.
- Geyer, P.E., Voytik, E., Treit, P.V., Doll, S., Kleinhempel, A., Niu, L., Müller, J.B., Buchholtz, M., Bader, J.M., Teupser, D., Holdt, L.M., and Mann, M. (2019) Plasma proteome profiling to detect and avoid sample-related biases in biomarker studies, *EMBO Mol. Med.*, doi: 10.15252/emmm.201910427.
- Banfálvi, G. (2011) Overview of cell synchronization, *Methods Mol. Biol.*, **761**, 1–23, doi: 10.1007/978-1-61779-182-6\_1.
- Emmert-Buck, M.R., Bonner, R.F., Smith, P.D., Chuaqui, R.F., Zhuang, Z., Goldstein, S.R., Weiss, R.A., and Liotta, L.A. (1996) Laser capture microdissection, *Science*, **274**, 998–1001, doi: 10.1126/science.274.5289.998.
- Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., and Enard, W. (2017) Comparative analysis of single-cell RNA sequencing methods, *Mol. Cell*, **65**, 631–643, doi: 10.1016/j.molcel.2017.01.023.
- Lee, J.H., Daugharthy, E.R., Scheiman, J., Kalhor, R., Yang, J.L., Ferrante, T.C., Terry, R., Jeanty, S.S.F., Li, C., Amamoto, R., Peters, D.T., Turczyk, B.M., Marblestone, A.H., Inverso, S.A., Bernard, A., Mali, P., Rios, X., Aach, J., and Church, G.M. (2014) Highly multiplexed subcellular RNA sequencing *in situ*, *Science*, **343**, 1360–1363, doi: 10.1126/science.1250212.

7. Lee, J.H., Daugharthy, E.R., Scheiman, J., Kalhor, R., Ferrante, T.C., Terry, R., Turczyk, B.M., Yang, J.L., Lee, H.S., Aach, J., Zhang, K., and Church, G.M. (2015) Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues, *Nat. Protoc.*, **10**, 442–458, doi: 10.1038/nprot.2014.191.
8. Picelli, S., Faridani, O.R., Björklund, Å.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014) Full-length RNA-seq from single cells using Smart-seq2, *Nat. Protoc.*, **9**, 171–181, doi: 10.1038/nprot.2014.006.
9. Valíhrach, L., Androvic, P., and Kubista, M. (2018) Platforms for single-cell collection and analysis, *Int. J. Mol. Sci.*, **19**, 807, doi: 10.3390/ijms19030807.
10. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., Gregory, M.T., Shuga, J., Montesclaros, L., Underwood, J.G., Masquelier, D.A., Nishimura, S.Y., Schnall-Levin, M., Wyatt, P.W., Hindson, C.M., Bharadwaj, R., Wong, A., Ness, K.D., Beppu, L.W., Deeg, H.J., McFarland, C., Loeb, K.R., Valente, W.J., Ericson, N.G., Stevens, E.A., Radich, J.P., Mikkelsen, T.S., Hindson, B.J., and Bielas, J.H. (2017) Massively parallel digital transcriptional profiling of single cells, *Nat. Commun.*, **8**, 14049, doi: 10.1038/ncomms14049.
11. Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lönnberg, P., and Linnarsson, S. (2014) Quantitative single-cell RNA-seq with unique molecular identifiers, *Nat. Methods*, **11**, 163–166, doi: 10.1038/nmeth.2772.
12. Zhang, X., Li, T., Liu, F., Chen, Y., Yao, J., Li, Z., Huang, Y., and Wang, J. (2019) Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems, *Mol. Cell*, **73**, 130–142, doi: 10.1016/j.molcel.2018.10.020.
13. Soldatov, R., Kaucka, M., Kastriti, M.E., Petersen, J., Chontorotzea, T., Englmaier, L., Akkuratova, N., Yang, Y., Häring, M., Dyachuk, V., Bock, C., Farlik, M., Piacentino, M.L., Boismoreau, F., Hilscher, M.M., Yokota, C., Qian, X., Nilsson, M., Bronner, M.E., Croci, L., Hsiao, W.-Y., Guertin, D.A., Brunet, J.-F., Consalez, G.G., Ernfors, P., Fried, K., Kharchenko, P.V., and Adameyko, I. (2019) Spatiotemporal structure of cell fate decisions in murine neural crest, *Science*, **364**, 9536, doi: 10.1126/science.aas9536.
14. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnberg, P., Furlan, A., Fan, J., Borm, L.E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundström, E., Castelo-Branco, G., Cramer, P., Adameyko, I., Linnarsson, S., and Kharchenko, P.V. (2018) RNA velocity of single cells, *Nature*, **560**, 494–498, doi: 10.1038/s41586-018-0414-6.
15. Burgess, D.J. (2018) Full speed ahead for single-cell analysis, *Nat. Rev. Genet.*, **19**, 668–669, doi: 10.1038/s41576-018-0049-3.
16. Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close, J.L., Long, B., Johansen, N., Penn, O., Yao, Z., Eggermont, J., Höllt, T., Levi, B.P., Shehata, S.I., Aevermann, B., Beller, A., Bertagnolli, D., Bruner, K., Casper, T., Cobbs, C., Dalley, R., Dee, N., Ding, S.-L., Ellenbogen, R.G., Fong, O., Garren, E., Goldy, J., Gwinn, R.P., Hirschstein, D., Keene, C.D., Keshk, M., Ko, A.L., Lathia, K., Mahfouz, A., Maltzer, Z., McGraw, M., Nguyen, T.N., Nyhus, J., Ojemann, J.G., Oldre, A., Parry, S., Reynolds, S., Rimorin, C., Shapovalova, N. V., Somasundaram, S., Szafer, A., Thomsen, E.R., Tieu, M., Quon, G., Scheuermann, R.H., Yuste, R., Sunkin, S.M., Lelieveldt, B., Feng, D., Ng, L., Bernard, A., Hawrylycz, M., Phillips, J.W., Tasic, B., Zeng, H., Jones, A.R., Koch, C., and Lein, E.S. (2019) Conserved cell types with divergent features in human versus mouse cortex, *Nature*, **573**, 61–68, doi: 10.1038/s41586-019-1506-7.
17. Khrameeva, E., Kurochkin, I., Han, D., Guijarro, P., Kanton, S., Santel, M., Qian, Z., Rong, S., Mazin, P., Bulat, M., Efimova, O., Tkachev, A., Guo, S., Sherwood, C.C., Camp, J.G., Paabo, S., Treutlein, B., and Khaitovich, P. (2019) Single-cell-resolution transcriptome map of human, chimpanzee, bonobo, and macaque brains, *bioRxiv*, **764936**, doi: 10.1101/764936.
18. Shekhar, K., and Menon, V. (2019) Identification of cell types from single-cell transcriptomic data, *Methods Mol. Biol.*, **1935**, 45–77, doi: 10.1007/978-1-4939-9057-3\_4.
19. Archakov, A., Ivanov, Y., Lisitsa, A., and Zgoda, V. (2009) Biospecific irreversible fishing coupled with atomic force microscopy for detection of extremely low-abundant proteins, *Proteomics*, **9**, 1326–1343, doi: 10.1002/pmic.200800598.
20. Aymoz, D., Wosika, V., Durandau, E., and Pelet, S. (2016) Real-time quantification of protein expression at the single-cell level via dynamic protein synthesis translocation reporters, *Nat. Commun.*, **7**, 11304, doi: 10.1038/ncomms11304.
21. Fulwyler, M.J. (1965) Electronic separation of biological cells by volume, *Science*, **150**, 910–911, doi: 10.1126/science.150.3698.910.
22. Picot, J., Guerin, C.L., Le Van Kim, C., and Boulanger, C.M. (2012) Flow cytometry: retrospective, fundamentals and recent instrumentation, *Cytotechnology*, **64**, 109–130, doi: 10.1007/s10616-011-9415-0.
23. Hughes, A.J., Spelke, D.P., Xu, Z., Kang, C.-C., Schaffer, D.V., and Herr, A.E. (2014) Single-cell western blotting, *Nat. Methods*, **11**, 749–755, doi: 10.1038/nmeth.2992.
24. Bendall, S.C., Simonds, E.F., Qiu, P., Amir, El-ad D., Krutzik, P.O., Finck, R., Bruggner, R. V., Melamed, R., Trejo, A., Ornatsky, O.I., Balderas, R.S., Plevritis, S.K., Sachs, K., Pe'er, D., Tanner, S.D., and Nolan, G.P. (2011) Single-cell mass cytometry of differential immune and drug responses across human hematopoietic continuum, *Science*, **332**, 687–696, doi: 10.1126/science.1198704.
25. Pali, C.G., Cheng, Q., Gillespie, M.A., Shannon, P., Mazurczyk, M., Napolitani, G., Price, N.D., Ranish, J.A., Morrissey, E., Higgs, D.R., and Brand, M. (2019) Single-cell proteomics reveal that quantitative changes in co-expressed lineage-specific transcription factors determine cell fate, *Cell Stem Cell*, **24**, 812–820, doi: 10.1016/j.stem.2019.02.006.
26. Marcon, E., Jain, H., Bhattacharya, A., Guo, H., Phanse, S., Pu, S., Byram, G., Collins, B.C., Dowdell, E., Fenner, M., Guo, X., Hutchinson, A., Kennedy, J.J., Krastins, B., Larsen, B., Lin, Z.-Y., Lopez, M.F., Loppnau, P., Miersch, S., Nguyen, T., Olsen, J.B., Paduch, M., Ravichandran, M., Seitova, A., Vadali, G., Vogelsang, M.S., Whiteaker, J.R., Zhong, G., Zhong, N., Zhao, L., Aebersold, R., Arrowsmith, C.H., Emili, A., Frappier, L., Gingras, A.-C., Gstaiger, M., Paulovich, A.G., Koide, S., Kossiakoff, A.A., Sidhu, S.S., Wodak, S.J., Gräslund, S., Greenblatt, J.F., and Edwards, A.M. (2015) Assessment of a method to characterize antibody selectivity and specificity for use in immunoprecipitation, *Nat. Methods*, **12**, 725–731, doi: 10.1038/nmeth.3472.
27. Coscia, F., Watters, K.M., Curtis, M., Eckert, M.A., Chiang, C.Y., Tyanova, S., Montag, A., Lastra, R.R.,

- Lengyel, E., and Mann, M. (2016) Integrative proteomic profiling of ovarian cancer cell lines reveals precursor cell associated proteins and functional status, *Nat. Commun.*, **7**, 12645, doi: 10.1038/ncomms12645.
28. Kaur, P., and O'Connor, P.B. (2007) Quantitative determination of isotope ratios from experimental isotopic distributions, *Anal. Chem.*, **79**, 1198–1204, doi: 10.1021/ac061535z.
  29. Ho, B., Baryshnikova, A., and Brown, G.W. (2018) Unification of protein abundance datasets yields a quantitative *Saccharomyces cerevisiae* proteome, *Cell Syst.*, **6**, 192–205, doi: 10.1016/j.cels.2017.12.004.
  30. Siwiak, M., and Zielenkiewicz, P. (2013) Transimulation – protein biosynthesis web service, *PLoS One*, **8**, e73943, doi: 10.1371/journal.pone.0073943.
  31. Virant-Klun, I., Leicht, S., Hughes, C., and Krijgsveld, J. (2016) Identification of maturation-specific proteins by single-cell proteomics of human oocytes, *Mol. Cell. Proteomics*, **15**, 2616–2627, doi: 10.1074/mcp.M115.056887.
  32. Sun, L., Dubiak, K.M., Peuchen, E.H., Zhang, Z., Zhu, G., Huber, P.W., and Dovichi, N.J. (2016) Single cell proteomics using frog (*Xenopus laevis*) blastomeres isolated from early stage embryos, which form a geometric progression in protein content, *Anal. Chem.*, **88**, 6653–6659, doi: 10.1021/acs.analchem.6b01921.
  33. Moroz, L.L. (2018) Neurosystematics and periodic system of neurons: model vs reference species at single-cell resolution, *ACS Chem. Neurosci.*, **9**, 1884–1903, doi: 10.1021/acschemneuro.8b00100.
  34. Chesnokova, E., Zuzina, A., Bal, N., Vinarskaya, A., Roshchin, M., Artyuhov, A., Dashinimaev, E., Aseyev, N., Balaban, P., and Kolosov, P. (2019) Experiments with snails add to our knowledge about the role of aPKC subfamily kinases in learning, *Int. J. Mol. Sci.*, **20**, 2117, doi: 10.3390/ijms20092117.
  35. Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A.K.A., and Hamon, C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS, *Anal. Chem.*, **75**, 1895–904, doi: 10.1021/ac0262560.
  36. Budnik, B., Levy, E., Harmange, G., and Slavov, N. (2018) SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation, *Genome Biol.*, **19**, 161, doi: 10.1186/s13059-018-1547-5.
  37. Huffman, R.G., Chen, A., Specht, H., and Slavov, N. (2019) DO-MS: data-driven optimization of mass spectrometry methods. *J. Proteome Res.*, **18**, 2493–2500, doi: 10.1021/acs.jproteome.9b00039.
  38. Chen, A.T., Franks, A., and Slavov, N. (2019) DART-ID increases single-cell proteome coverage, *PLOS Comput. Biol.*, **15**, e1007082, doi: 10.1371/journal.pcbi.1007082.
  39. Specht, H., Emmott, E., Perlman, D.H., Koller, A., and Slavov, N. (2019) High-throughput single-cell proteomics quantifies the emergence of macrophage heterogeneity, *bioRxiv*, 665307, doi: 10.1101/665307.
  40. Dou, M., Clair, G., Tsai, C.-F., Xu, K., Chrisler, W.B., Sontag, R.L., Zhao, R., Moore, R.J., Liu, T., Pasa-Tolic, L., Smith, R.D., Shi, T., Adkins, J.N., Qian, W.-J., Kelly, R.T., Ansong, C., and Zhu, Y. (2019) High-throughput single cell proteomics enabled by multiplex isobaric labeling in a nanodroplet sample preparation platform, *Anal. Chem.*, 9b03349, doi: 10.1021/acs.analchem.9b03349.
  41. Zhu, Y., Piehowski, P.D., Zhao, R., Chen, J., Shen, Y., Moore, R.J., Shukla, A.K., Petyuk, V.A., Campbell-Thompson, M., Mathews, C.E., Smith, R.D., Qian, W.-J., and Kelly, R.T. (2018) Nanodroplet processing platform for deep and quantitative proteome profiling of 10–100 mammalian cells, *Nat. Commun.*, **9**, 882, doi: 10.1038/s41467-018-03367-w.
  42. Schoof, E.M., Rapin, N., Savickas, S., Gentil, C., Lechman, E., Haile, J.S., auf dem Keller, U., Dick, J.E., and Porse, B.T. (2019) A quantitative single-cell proteomics approach to characterize an acute myeloid leukemia hierarchy, *bioRxiv*, 745679, doi: 10.1101/745679.
  43. Johansson, H.J., Socciarelli, F., Vacanti, N.M., Haugen, M.H., Zhu, Y., Siavelis, I., Fernandez-Woodbridge, A., Aure, M.R., Sennblad, B., Vesterlund, M., Branca, R.M., Orre, L.M., Huss, M., Fredlund, E., Beraki, E., Garred, Ø., Boekel, J., Sauer, T., Zhao, W., Nord, S., Högländer, E.K., Jans, D.C., Brismar, H., Haukaas, T.H., Bathen, T.F., Schlichting, E., Naume, B., Luders, T., Borgen, E., Kristensen, V.N., Russnes, H.G., Lingjaerde, O.C., Mills, G.B., Sahlberg, K.K., Børresen-Dale, A.-L., and Lehtiö, J. (2019) Breast cancer quantitative proteome and proteogenomic landscape, *Nat. Commun.*, **10**, 1600, doi: 10.1038/s41467-019-09018-y.
  44. Dimitrakopoulos, L., Prassas, I., Diamandis, E.P., Nesvizhskii, A., Kislinger, T., Jaffe, J., and Drabovich, A. (2016) Proteogenomics: opportunities and caveats, *Clin. Chem.*, **62**, 551–557, doi: 10.1373/clinchem.2015.247858.
  45. Smith, L.M., and Kelleher, N.L. (2013) Proteoform: a single term describing protein complexity. *Nat. Methods*, **10**, 186–187, doi: 10.1038/nmeth.2369.
  46. Simões, A.E., Pereira, D.M., Amaral, J.D., Nunes, A.F., Gomes, S.E., Rodrigues, P.M., Lo, A.C., D'Hooge, R., Steer, C.J., Thibodeau, S.N., Borralho, P.M., and Rodrigues, C.M. (2013) Efficient recovery of proteins from multiple source samples after trizol(®) or trizol(®)LS RNA extraction and long-term storage, *BMC Genomics*, **14**, 181, doi: 10.1186/1471-2164-14-181.
  47. Mun, D.-G., Bhin, J., Kim, S., Kim, H., Jung, J.H., Jung, Y., Jang, Y.E., Park, J.M., Kim, H., Jung, Y., Lee, H., Bae, J., Back, S., Kim, S.-J., Kim, J., Park, H., Li, H., Hwang, K.-B., Park, Y.S., Yook, J.H., Kim, B.S., Kwon, S.Y., Ryu, S.W., Park, D.Y., Jeon, T.Y., Kim, D.H., Lee, J.-H., Han, S.-U., Song, K.S., Park, D., Park, J.W., Rodriguez, H., Kim, J., Lee, H., Kim, K.P., Yang, E.G., Kim, H.K., Paek, E., Lee, S., Lee, S.-W., and Hwang, D. (2019) Proteogenomic characterization of human early-onset gastric cancer, *Cancer Cell*, **35**, 111–124, doi: 10.1016/j.ccell.2018.12.003.
  48. Poirion, O., Zhu, X., Ching, T., and Garmire, L.X. (2018) Using single nucleotide variations in single-cell RNA-seq to identify subpopulations and genotype-phenotype linkage, *Nat. Commun.*, **9**, 4892, doi: 10.1038/s41467-018-07170-5.
  49. Levitsky, L.I., Kliuchnikova, A.A., Kuznetsova, K.G., Karpov, D.S., Ivanov, M.V., Pyatnitskiy, M.A., Kalinina, O.V., Gorshkov, M.V., and Moshkovskii, S.A. (2019) Adenosine-to-inosine RNA editing in mouse and human brain proteomes, *Proteomics*, 1900195, doi: 10.1002/pmic.201900195.
  50. Ximerakis, M., Lipnick, S.L., Innes, B.T., Simmons, S.K., Adiconis, X., Dionne, D., Mayweather, B.A., Nguyen, L., Niziolik, Z., Ozek, C., Butty, V.L., Isserlin, R., Buchanan, S.M., Levine, S.S., Regev, A., Bader, G.D., Levin, J.Z., and Rubin, L.L. (2019) Single-cell transcriptomic profiling of the aging mouse brain, *Nat. Neurosci.*, **22**, 1696–1708, doi: 10.1038/s41593-019-0491-3.



## SINGLE CELL PROTEOGENOMICS AS AN IMMEDIATE PROSPECT

## Review

S. A. Moshkovskii<sup>1,2\*</sup>, A. A. Lobas<sup>3</sup>, and M. V. Gorshkov<sup>3</sup><sup>1</sup> Pirogov Russian National Research Medical University, 117997 Moscow, Russia; E-mail: smosh@mail.ru<sup>2</sup> Orekhovich Institute of Biomedical Chemistry, 119121 Moscow, Russia<sup>3</sup> Talroze Institute for Energy Problems of Chemical Physics, Semenov Federal Research Center of Chemical Physics, Russian Academy of Sciences, 119334 Moscow, Russia

Received October 15, 2019

Revised November 11, 2019

Accepted November 11, 2019

Recently technical advances in genomic technology led to explosive growth of transcriptome-wide studies at the level of single cells. The review describes the beginning of single-cell proteomics which started soon after transcriptomic methods were developed. Thus, first studies have been published that used liquid chromatography-mass spectrometry to analyze shotgun proteomes of single cells. In these works, cells were separated by methods used in transcriptomics, e.g., by cell sorting, and mass-spectrometric analysis was performed by a modified method of tandem mass tags. Data integration of single cell transcriptomics and proteomics as a proteogenomic approach will provide better understanding of mechanisms of cell interactions in normal development and disease.

*Keywords:* proteomics, transcriptomics, proteogenomics, single cell analysis, target mass tag (TMT), mass spectrometry